

**UNIVERSIDAD AUTÓNOMA DE MADRID**

**ESCUELA POLITÉCNICA SUPERIOR**



**TRABAJO FIN DE GRADO**

**Modelado y detección de elementos de  
interés en secuencias de vídeo de carreteras  
mediante técnicas de visión artificial**

Grado en Ingeniería de Tecnologías y Servicios de  
Telecomunicación

**Alberto Mozo Robles**  
**Tutor: Luis Salgado Álvarez de Sotomayor**

**Julio 2014**



# Resumen y palabras clave

## Resumen

Los sistemas avanzados de ayuda a la conducción (ADAS, *Advanced Driver Assistance System*) están teniendo un importante crecimiento y desarrollo en los últimos años, aumentando en gran medida la seguridad en la conducción. Entre estos sistemas destacan los basados en vídeo, cuyo funcionamiento reside en el análisis de imágenes de la escena.

Dichos sistemas necesitan información fiable y robusta del plano de la carretera, destacando la detección de líneas divisorias de carretera, para facilitar y mejorar la tarea de análisis posterior.

Este trabajo de final de grado tiene como objetivo la implementación de métodos que consigan obtener esa información fiable, eliminando la distorsión de perspectiva que presentan las imágenes capturas por la cámara. Se trata de obtener una imagen rectificadora del plano de carretera, obteniendo una vista cenital.

Se han desarrollado dos implementaciones diferentes para la consecución del objetivo, que difieren en la cantidad y tipo de información previa requerida sobre la escena.

Se van a comparar ambas alternativas, así como el efecto de considerar el cálculo del punto de fuga de forma robusta para conseguir una estabilización y mejora en la imagen rectificadora.

## Palabras clave

Rectificación de plano, homografía, matriz de proyección, calibración, punto de fuga, Transformada de Hough, IPM (*Inverse Perspective Mapping*), vista cenital, vista de pájaro (*bird-view*), estabilización.

# Abstract and Keywords

## Abstract

Advanced Driver Assistance System (ADAS) are having a significant growth and development in recent years, greatly increasing driving safety. Amongst these, we must highlight the video based systems whose operation lies in the analysis of images of the scene.

Such systems need a reliable and robust road plane, emphasizing the detection of dividing lines of the road to facilitate and enhance further analysis tasks.

This final degree work aims the implementation of methods to succeed in obtaining such reliable information, eliminating the perspective distortion shown by the images captured by the camera. It is about getting a rectified image of the road plane, getting an bird-view.

We have developed two different implementations to achieve the objective, which differ in the amount and type of prior information required on the scene.

Both alternatives are compared, as well as the effect of computing robustly the vanishing point to achieve stabilized and improved rectified images.

## Keywords

Plane rectification, homography, projection matrix, calibration, vanishing point, Hough Transform, IPM (Inverse Perspective Mapping), bird's eye view, stabilization.



# Agradecimientos

*Antes de comenzar, no podía olvidarme de mencionar a aquellas personas que han hecho posible la realización de este trabajo.*

*En primer lugar, agradecer la oportunidad que me ha brindado mi tutor, Luis Salgado, de llevar a cabo este proyecto. Es de agradecer su paciencia, su comprensión y toda la ayuda que me ha ofrecido en todo momento.*

*Igual de importante han sido mis compañeros de clase estos cuatro años. Parece que no llegaríamos nunca a este momento y, una vez aquí, nos damos cuenta de lo rápido que ha pasado el tiempo. Por su ayuda, las risas y los buenos momentos tanto dentro como fuera de las aulas.*

*También quiero mencionar a mis amigos de siempre, con los que he pasado tantos y tan buenos momentos; esos veranos que consiguen darte fuerzas para seguir cada año dando lo máximo.*

*Por supuesto, me gustaría agradecer a aquellas personas que han estado apoyándome y animándome estos últimos meses, soportando mis pesados comentarios sobre el TFG día tras día y, sobretudo, mostrando una enorme confianza en mí que me hacía sentir capaz de todo.*

*Por último, mis agradecimientos más especiales y sinceros son, como no podía ser de otra manera, para mi familia. Sin ellos, nunca habría llegado a donde estoy. Por su ayuda, su confianza ciega en mí, su apoyo continuo y motivación. Por todo, este trabajo también es vuestro.*

*Muchas gracias.*

*Alberto Mozo Robles.  
Julio 2014.*



# Índice general

Índice figuras.....	VIII
Índice tablas.....	IX
Glosario .....	XIII
<b>1. Introducción .....</b>	<b>1</b>
1.1 Motivación.....	1
1.2 Objetivos.....	1
1.3 Organización de la memoria.....	2
<b>2. Estado del Arte .....</b>	<b>3</b>
2.1 Introducción a los sistemas de ayuda a la conducción .....	3
2.2 Sistemas de Ayuda a la Conducción basados en vídeo .....	3
2.3 Fundamentos Teóricos .....	4
2.3.1 Geometría Proyectiva .....	4
2.3.2 Homografía .....	7
2.4 Casos de rectificación del plano de carretera .....	9
2.4.1 Matriz de proyección conocida .....	9
2.4.2 Parámetros extrínsecos desconocidos .....	10
2.4.2.1 Ángulo de alabeo desconocido .....	10
2.4.2.2 Rotación desconocida .....	11
2.4.3 Parámetros extrínsecos e intrínsecos desconocidos .....	11
2.4.3.1 Sin calcular la matriz de proyección .....	11
<b>3. Diseño .....</b>	<b>13</b>
3.1 Estrategia general .....	13
3.2 Extracción de características .....	15
3.2.1 Aplicación detector de bordes .....	15
3.2.2 Filtrado adaptado .....	16
3.3 Determinación de líneas divisorias y punto de fuga .....	19
3.3.1 Transformada de Hough .....	19
3.3.2 Cálculo de máximos locales .....	22
3.3.3 Ajuste de los máximos de la Transformada de Hough a una senoide .....	23
3.3.4 Filtrado del punto de fuga .....	25
3.4 Rectificación del Plano.....	27
3.4.1 Cómputo de la Homografía mediante correspondencia entre puntos .....	27
3.4.2 Cálculo de la IPM mediante la Matriz de Proyección .....	32
<b>4. Resultados.....</b>	<b>35</b>
4.1 Impacto en el punto de fuga .....	36
4.1.1 Parámetros de la extracción de características.....	36
4.1.2 Parámetros de la determinación de líneas de carretera y punto de fuga ..	37
4.2 Impacto en la construcción de la IPM .....	40



4.2.1 Parámetros del Método 1 .....	40
4.2.2 Parámetros del Método 2 .....	41
4.3 Comparación ambos métodos.....	42
<b>5. Conclusiones y Trabajos Futuros .....</b>	<b>45</b>
5.1 Conclusiones.....	45
5.2 Trabajos futuros.....	46
<b>Referencias .....</b>	<b>47</b>
<b>Anexos .....</b>	<b>LI</b>
A.1 Algoritmo DLT.....	LI
A.2 Tablas.....	LIII

# Índice figuras

Figura 1.1: Rectificación Plano Carretera: (a) Imagen original; (b) IPM ( <i>Inverse Perspective Mapping</i> ) .....	2
Figura 2.1: Modelo de Cámara Puntual ( <i>Pinhole Camera</i> ) .....	4
Figura 2.2: Relación sistema de coordenadas de la cámara y el mundo 3D. ....	6
Figura 2.3: Ángulos de rotación parámetros extrínsecos: (a) Cabeceo; (b) Guiñada; (c) Alabeo [26] .....	7
Figura 2.4: Relación entre dos imágenes de un mismo plano: Homografía. ....	8
Figura 2.5: Explicación Homografía .....	8
Figura 2.6: Detección segundo punto de fuga [26]. ....	11
Figura 2.7: Computación de la homografía entre imágenes mediante correspondencia de cuatro puntos [26]. ....	12
Figura 3.1: Generación IPM: (a) Imagen original; (b) IPM ( <i>Inverse Perspective Mapping</i> ) .....	13
Figura 3.2: Detección de bordes con el operador Canny. ....	16
Figura 3.3: Filtrado líneas carretera: (a) Imagen original; (b) Imagen filtrada, $\tau=10$ ; (c) Imagen filtrada, $\tau=5$ ; (d) Imagen filtrada, $\tau=40$ .....	17
Figura 3.4: Efecto del filtrado para una fila. ....	17
Figura 3.5: Histograma imagen bimodal. ....	18
Figura 3.6: Búsqueda iterativa del umbral que equidista de las medias superior e inferior. ....	18
Figura 3.7: Correspondencia Espacio – Espacio transformado Hough [27]. ....	19
Figura 3.8: Acumulador plano de Hough.....	20
Figura 3.9: Correspondencia Espacio – Espacio transformado Hough [27]. ....	21
Figura 3.10: (a) Transformada de Hough; (b) Dilatación Transformada de Hough; (c) Detección de máximos en la Transformada de Hough. ....	21
Figura 3.11: Aplicación Operador morfológico dilatación: (a) Imagen original; (b) Imagen dilatada .....	22

Figura 3.12: Resultado de la Transformada de Hough .....	23
Figura 3.13: Ajuste a senoide: (a) Senoide en la Transformada de Hough; (b) Punto de fuga calculado a partir de la senoide. ....	24
Figura 3.14: Mal ajuste a senoide: (a) Senoide en la Transformada de Hough; (b) Punto de fuga calculado a partir de la senoide. ....	24
Figura 3.15: Ajuste a una senoide filtrado: (a) Senoide en la Transformada de Hough; (b) Punto de fuga calculado a partir de la senoide. ....	25
Figura 3.16: Variación punto de fuga y punto de fuga filtrado: (a) Variación eje-x punto de fuga; (b) Variación eje-x punto de fuga filtrado; (c) Variación eje-y punto de fuga; (d) Variación eje-y punto de fuga filtrado. ....	26
Figura 3.17: Representación del punto de fuga (línea continua) y punto de fuga filtrado (línea discontinua). ....	26
Figura 3.18: Reestimación rectas (líneas rojas) a partir del punto de fuga (azul).....	28
Figura 3.19: Selección de puntos de correspondencia: (a) Puntos formando un trapecio en la imagen original; (b) Puntos formando un cuadrado en la IPM .....	29
Figura 3.20: Selección de puntos de correspondencia en IPM: (a) Poca separación en ambos ejes; (b) Separación grande en el eje-y; (c) Separación grande en el eje-x; (d) Separación grande en ambos ejes.....	29
Figura 3.21: Homografía ( $H$ ) y Homografía inversa ( $H^{-1}$ ).....	31
Figura 3.22: Interpolación bilineal.....	31
Figura 3.23: Imagen Rectificado mediante cuatro Puntos de Correspondencias: (a) Redondeo; (b) Interpolación bilineal.....	32
Figura 3.24: IPM sin estabilización .....	33
Figura 3.25: Comparativa IPM estabilizada (a) y IPM No estabilizada (b) .....	34
Figura 4.1: Verificación de la bondad del método 2 a partir de una secuencia sintética (a).....	35
Figura 4.2: Comparación punto de fuga calculado (azul) y punto de fuga teórico (rojo) .....	36
Figura 4.3: Comparación punto de fuga calculado (azul) y punto de fuga teórico (rojo) .....	36

Figura 4.4: Líneas detectadas tras Filtro adaptado .....	37
Figura 4.5: Comparación punto de fuga calculado (azul) y punto de fuga teórico (rojo) .....	37
Figura 4.6: Líneas detectas con número de máximos elevado .....	38
Figura 4.7: Comparativa entre ajuste filtrado (b) y normal (a) .....	38
Figura 4.8: Evolución punto fuga calculado (azul) vs punto de fuga teórico (rojo): (a) ajuste normal; (b): ajuste filtrado.....	38
Figura 4.9: Comparación punto fuga calculado (verde) y filtrado (azul): (a) Filtrado con ; (b) Filtrado con ... ..	39
Figura 4.10: Filtrado del punto de fuga: verde: punto de fuga calculado; azul: punto de fuga filtrado; rojo: punto de fuga teórico. ....	39
Figura 4.11: Gestión oclusiones simple: (a) Detección rectas en la imagen original; (b): IPMs generadas incorrectamente .....	41
Figura 4.12: Gestión oclusiones avanzada: (a) Detección rectas en la imagen original; (b) IPMs generadas correctamente .....	41
Figura 4.13: Comparativa diferentes profundidades .....	42
Figura 4.14: Comparativa de IPM no estabilizada (a) y estabilizada (b). ....	42
Figura 4.15: Generación IPM en línea recta: (a) Imagen original; (b) IPM Método 1; (c) IPM Método 2.....	43
Figura 4.16: Generación IPM en cambio de carril: (a) Imagen original; (b) IPM Método 1; (c) IPM Método 2.....	43

# Índice tablas

Tabla A.2.1: Valores teórico-calculado del punto de fuga para Canny, mejor situación ..... LIII

Tabla A.2.2: Valores teórico-calculado del punto de fuga para Canny, mala situación. LIII

Tabla A.2.3: Valores teórico-calculado del punto de fuga para Filtrado adaptado ..... LIV



# Glosario

**TIC:** Tecnologías de la Información y la Comunicación

**ADAS:** Advanced Driver Assistance System

**Rectificación del plano:** transformación de un plano de imagen en otro que no presenta distorsión de perspectiva

**IPM:** Inverse Perspective Mapping

**Vista cenital (*bird-view*):** punto de vista con la cámara situada perpendicular respecto del suelo, ofreciendo un campo de visión de arriba a abajo.

**Líneas divisorias de carril:** líneas blancas que delimitan los carriles de la carretera.

**Matriz de proyección:** determina la proyección de una escena 3D en una imagen 2D.

**Punto de fuga:** lugar geométrico donde las proyecciones de rectas paralelas en el espacio convergen.

**Homografía:** transformación proyectiva que determina una correspondencia entre dos figuras geométricas planas.

**Transformada de Hough:** es un sistema de votaciones empleado para la detección de figuras en imágenes digitales.

**Coordenadas homogéneas:** descripción de un punto bidimensional con tres coordenadas.

**Calibración cámara:** estimación de los parámetros extrínsecos e intrínsecos que definen las condiciones de formación de la imagen para una cámara.





# 1. Introducción:

## 1.1 Motivación:

Las Tecnologías de la Información y la Comunicación (TIC) son en la actualidad una gran ayuda y una poderosa herramienta aplicable en prácticamente todos los ámbitos de la vida humana.

Diferentes técnicas de visión artificial están creciendo con fuerza en el ámbito automovilístico. En él encontramos cada vez más numerosos y novedosos sistemas de ayuda a la conducción, conocidos como ADAS (*Advanced Driver Assistance Systems*). Son sistemas que buscan ayudar en el proceso de la conducción, facilitando y mejorando la experiencia del conductor al volante, con el único objetivo de incrementar la seguridad en la carretera.

Entre ellos destacan los sistemas basados en vídeo mono-cámara, debido a su buen rendimiento con bajo coste. Abarcan desde sistemas que controlan la velocidad, que avisan cuando el vehículo se está yendo a otro carril, que evitan colisiones, hasta aquellos que son capaces de reconocer las señales de tráfico. Todos estos sistemas hacen uso de secuencias de vídeo de carreteras, procesan las imágenes y obtienen la información necesaria para llevar a cabo su función.

Obtener la información presente en la escena (extracción de las características de la imagen) es una parte fundamental del proceso para su posterior análisis. Por ello, resulta de vital importancia conseguir una descripción del entorno fiable y robusta. Para este propósito aplicamos una rectificación del plano de carretera, consiguiendo una imagen que no presenta distorsión de perspectiva: IPM (*Inverse Perspective Mapping*), vista cenital o vista de pájaro (*bird-view*).

## 1.2 Objetivos:

El objetivo de este trabajo de fin de grado es el de realizar una rectificación del plano de carretera, con el fin de obtener una vista de la escena en la que ha sido eliminada la distorsión provocada por la perspectiva (líneas paralelas no lo son en la imagen, las distancias no son correctas, etc.). Esta nueva vista (Figura 1.1) recibe el nombre de vista cenital, vista de pájaro (*bird-view*) o IPM (*Inverse Perspective Mapping*).

Para la consecución de dicho objetivo es necesario conocer y/o extraer información de la escena. Se van a desarrollar dos alternativas, que difieren en las hipótesis de partida y en la información requerida a priori referente al sistema de adquisición:

La primera de ellas asume que la escena presentará líneas divisorias de carril y que éstas se encontrarán en un plano (el plano de carretera). Necesitará determinar la localización de las mismas para encontrar puntos de correspondencia.

La segunda asume la correcta colocación de la cámara en el vehículo durante el proceso de adquisición y empleará el conocimiento de la matriz de proyección para la rectificación.

En ambos casos, la obtención del punto de fuga dominante en la escena es necesario, por lo que otro de los objetivos será su cálculo de forma robusta.



Figura 1.1: Rectificación Plano Carretera: (a) Imagen original; (b) IPM (*Inverse Perspective Mapping*)

### 1.3 Organización de la memoria:

La memoria de este trabajo queda estructurada de la siguiente manera:

- Capítulo 1: Introducción y motivación del trabajo
- Capítulo 2: Estado del arte de los sistemas de ayuda a la conducción. Fundamentos teóricos sobre la geometría proyectiva. Explicación diferentes métodos de rectificación de planos.
- Capítulo 3: Diseño del sistema completo, formado por tres bloques: extracción de características; determinación de líneas de carretera y punto de fuga; cómputo de la IPM mediante dos alternativas.
- Capítulo 4: Evaluación del sistema: análisis de resultados obtenidos para las diferentes alternativas.
- Capítulo 5: Conclusiones obtenidas tras los resultados y propuestas de trabajos futuros para mejora del sistema.
- Referencias bibliográficas consultadas.
- Anexos con información en detalle.

## 2. Estado del Arte

### 2.1 Introducción a los sistemas de ayuda a la conducción:

Los sistemas avanzados de ayuda a la conducción (ADAS, *Advanced Driver Assistance Systems*) surgen con un único objetivo: reducir la gravedad de los accidentes y evitar los posibles impactos por culpa de la falta de atención al conductor.

Los datos de mortalidad en accidentes de tráfico de 2002 [5] impulsaron el buscar soluciones para reducir esa cifra. Se empezaron a explotar los avances de la tecnología para producir sistemas inteligentes de ayuda a la conducción con la finalidad de aumentar la seguridad vial.

### 2.2 Sistemas de Ayuda a la Conducción basados en vídeo:

Existen multitud de sistemas de ayuda a la conducción, siendo de especial importancia aquellos basados en vídeo [8]. Debido a la mejora computacional de los procesadores, hoy en día es posible emplear sensores de visión en estos procesos [3].

Los sistemas basados en vídeo, en comparación con otros como radar o lidar, obtienen más información relevante sobre la escena del plano de carretera, siendo una mejor alternativa para los sistemas de ayuda a la conducción [8].

Los sistemas de ayuda a la conducción precisan de ciertos requisitos: tienen que ejecutarse en tiempo real y obtener información de calidad, precisión en los resultados y robustez frente a diferentes situaciones. El entorno de trabajo es muy complejo debido a la gran variabilidad de la escena (cambios de iluminación, oclusiones de líneas de carretera, condiciones meteorológica, etc.)

Dentro de los sistemas basados en vídeo encontramos aquéllos que emplean sistemas de estereovisión, debido a su capacidad para recuperar la información de profundidad mediante el análisis de dos secuencias de vídeo sincronizadas. Muchos trabajos consiguen precisión sub-píxel en el modelado de las líneas de carretera [20-22]. Otros usan la alineación entre las imágenes estéreo para detectar objetos volumétricos en el plano de la carretera [23, 24].

A pesar de que estos sistemas son capaces de obtener la información de profundidad, los sistemas de visión estereoscópica tienen varios inconvenientes: los sistemas multi-vista no son considerados para aplicaciones de tiempo real, dada su complejidad: necesitan perfecta calibración, sistemas de adquisición sincronizados y existe gran dificultad en encontrar correspondencias fiables entre las imágenes.

Los sistemas mono-cámaras son más efectivos computacionalmente por lo que son muy utilizados en aplicaciones reales [25]. Dentro de este tipo de sistemas encontramos: ayudas para el aparcamiento, control de velocidad, sistemas de frenada

de emergencia, sistemas de reconocimiento de señales de tráfico, asistente de visión nocturna, aviso de salida de carril, detector de peatones, etc. Para el cumplimiento de los objetivos de dichos sistemas es necesario realizar tareas de detección de obstáculos, detección de vehículos [7, 10, 13], seguimiento de las líneas de la carretera [10, 11, 13], estimación de la velocidad del vehículo [9], estimación del movimiento y localización [3].

Para llevar a cabo estas tareas, los sistemas requieren información de las líneas divisorias de la carretera, curvatura de la misma, información del vehículo de delante, etc. Por ello, se coloca una cámara en el salpicadero del coche, capturando la vista frontal, generando una imagen en perspectiva de la carretera. El problema existente en esta imagen es la distorsión que presente en la forma de la carretera: las líneas no son paralelas, fugan en un punto, el punto de fuga; la anchura de las líneas y los carriles no es correcta; y las distancias no son reales.

Por este motivo, muchos sistemas basados en visión necesitan realizar una rectificación de la imagen [26] con el objetivo de eliminar dicha distorsión y conseguir así simplificar las etapas de análisis posteriores. Esta imagen rectificada recibe el nombre de IPM (*Inverse Perspective Mapping*).

## 2.3 Fundamentos Teóricos:

Para entender y comprender las necesidades y procedimientos que se llevan a cabo en el diseño de este trabajo, es necesario tener ciertas nociones sobre la geometría proyectiva y las homografías.

### 2.3.1 Geometría Proyectiva:

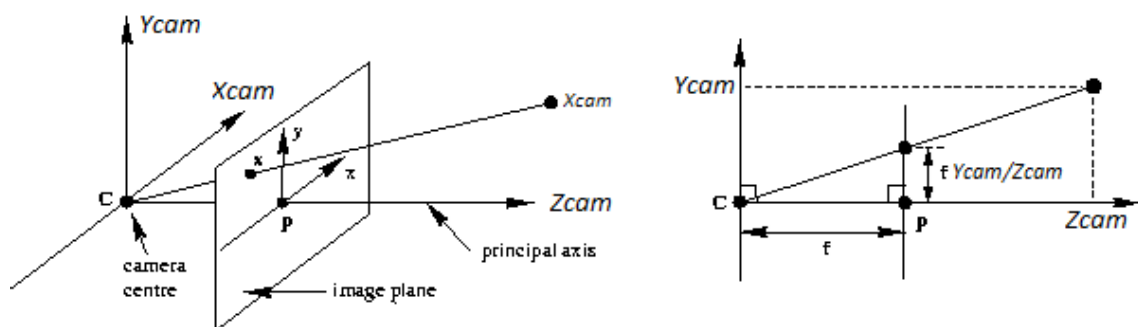


Figura 2.1: Modelo de Cámara Puntual (*Pinhole Camera*) [29].

La Figura 2.1 muestra el **modelo de cámara puntual** (*Pinhole Camera*). Los parámetros que gobiernan dicho modelo son:

- Plano imagen: plano de proyección
- Centro óptico o foco:  $C$ , *pinhole*

- Distancia focal:  $f$  , distancia entre el foco y el plano de la imagen
- Eje óptico o principal (*principal axis*)
- Punto principal:  $P$  , intersección entre el eje principal y el plano de la imagen. Suele estar cerca del centro de la imagen.
- Proyección  $x$  del punto 3D  $x_{cam}$  : intersección entre la recta  $\overline{x_{cam}C}$  y el plano de imagen.

El modelo de cámara puntual describe la relación matemática entre las coordenadas de un punto 3D y su proyección en el plano de la imagen (proyección de una escena 3D en una imagen 2D). La relación entre las coordenadas del punto 3D ( $x_{cam}$ ) y su proyección ( $x$ , con  $z = f$ ) es:

$$x = f \frac{x_{cam}}{z_{cam}}, \quad y = f \frac{y_{cam}}{z_{cam}}$$

$$\begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \approx \begin{pmatrix} fx_{cam} \\ fy_{cam} \\ z_{cam} \end{pmatrix} = \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{pmatrix} x_{cam} \\ y_{cam} \\ z_{cam} \\ 1 \end{pmatrix} \quad (2.1)$$

Con la fórmula anterior (2.1) conseguimos realizar la proyección de los puntos del mundo real, al plano de la imagen. No obstante, el modelo de cámara puntual incluye dos procesos más: cambio del sistema de referencia en el plano (parámetros intrínsecos) y el cambio del sistema de referencia en el espacio (parámetros extrínsecos).

**Parámetros intrínsecos:** son aquellos que definen la geometría interna y la óptica de la cámara. Son constantes siempre que no varíen las características y posiciones relativas entre la óptica y el sensor de la cámara. Vienen definidos en la matriz de proyección.

$K$  es una matriz triangular superior, 3x3, llamada matriz de calibración:

$$K = \begin{bmatrix} \alpha_x & & x_0 \\ & \alpha_y & y_0 \\ & & 1 \end{bmatrix},$$

donde  $\alpha_x$  y  $\alpha_y$  son el escalado de la imagen en las direcciones  $x$  e  $y$  respectivamente.

$(x_0, y_0)$  es el punto principal: intersección entre eje óptico y plano de la imagen.

$\alpha_y / \alpha_x$  es la relación de aspecto (*aspect ratio*).

La proyección de los puntos en coordenadas de la cámara a coordenadas de la imagen viene determinada por:

$$\begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = \frac{1}{f} \begin{bmatrix} \alpha_x & x_0 \\ \alpha_y & y_0 \\ 1 & f \end{bmatrix} \begin{pmatrix} x_{cam} \\ y_{cam} \\ f \end{pmatrix} = K \begin{pmatrix} x_{cam} \\ y_{cam} \\ f \end{pmatrix}, \quad (2.2)$$

**Parámetros extrínsecos:** son aquéllos que relacionan los sistemas de referencia del mundo real y la cámara, describiendo la posición y orientación de la cámara en el sistema de coordenadas del mundo real.

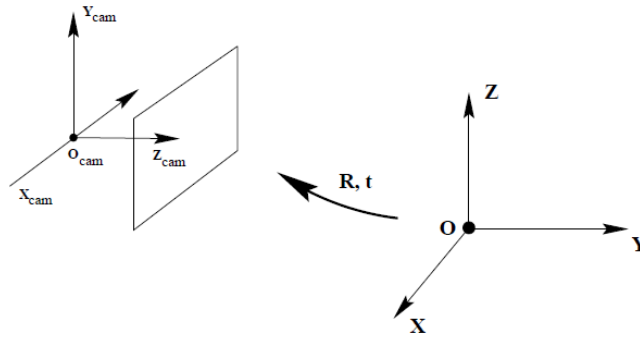


Figura 2.2: Relación sistema de coordenadas de la cámara y el mundo 3D. [29]

La transformación entre dichos sistemas de coordenadas viene dada por:

$$\begin{pmatrix} x_{cam} \\ y_{cam} \\ z_{cam} \\ 1 \end{pmatrix} = \begin{bmatrix} R & t \\ 0^T & 1 \end{bmatrix} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix}, \quad (2.3)$$

donde  $t$  es el vector de traslación  $3 \times 1$ .  $t = (t_x, t_y, t_z)^T$ . La altura de la cámara ( $t_y$ ) se considera un valor constante, mientras que  $t_z$  y  $t_x$  son fijados por definición, ya que el sistema de coordenadas de la cámara avanza conjuntamente con el sistema de coordenadas del mundo real.

$R$  es la matriz de rotación  $3 \times 3$ . Dicha matriz está compuesta por tres rotaciones diferentes:

$$R = R_\theta R_\gamma R_\beta = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \theta & \sin \theta \\ 0 & -\sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} \cos \gamma & 0 & -\sin \gamma \\ 0 & 1 & 0 \\ \sin \gamma & 0 & \cos \gamma \end{pmatrix} \begin{pmatrix} \cos \beta & \sin \beta & 0 \\ -\sin \beta & \cos \beta & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

donde  $\theta$ ,  $\gamma$  y  $\beta$  se corresponden respectivamente con los ángulos : cabeceo (*pitch*,  $\theta$ ), guiñada (*yaw*,  $\gamma$ ) y alabeo (*roll*,  $\beta$ ).

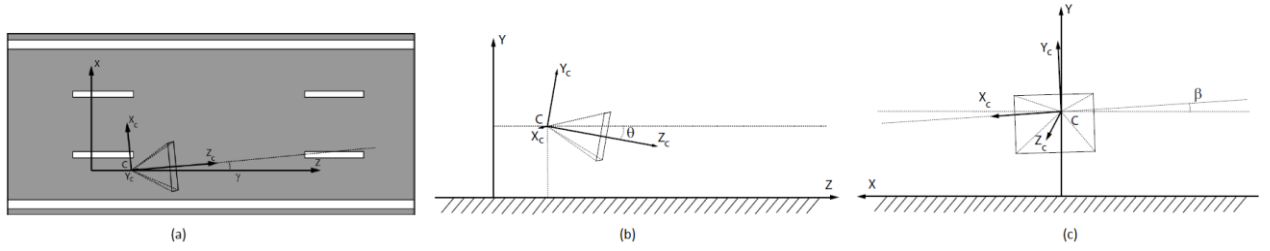


Figura 2.3: Ángulos de rotación parámetros extrínsecos: (a) Cabeceo; (b) Guiñada; (c) Alabeo [26]

- **Cabeceo (*pitch*)**: es el ángulo de rotación con respecto al eje X. Puede cambiar debido a pendientes y baches o saltos en la carretera.
- **Guiñada (*yaw*)**: es el ángulo de rotación con respecto al eje Y. Este ángulo depende de la instalación de la cámara. Normalmente es distinto de cero, puesto que la cámara se instala a la izquierda o la derecha del espejo retrovisor y se rota para mostrar el carril en el centro de la imagen. Está relacionado con el movimiento de la dirección de las ruedas.
- **Alabeo (*roll*)**: es el ángulo de rotación con respecto al eje Z. Si la cámara está correctamente instalada debería ser cero.

Concatenando las 3 matrices (proyección cónica (2.1), matriz calibración (2.2) y rotación-traslación (2.3)):

$$\begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = K \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} R & t \\ 0^T & 1 \end{bmatrix} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} = K[R | t]X \quad (2.4)$$

Se define la **matriz de proyección P** (3x4) que define la transformación de un espacio de 3 dimensiones a una imagen:

$$x = PX \quad P = K[R | t] = KR[I | R^T t] \quad (2.5)$$

### 2.3.2 Homografía:

El cambio de plano entre dos imágenes está gobernado por una Homografía. En la geometría, denominamos Homografía a toda transformación proyectiva que determina la correspondencia entre dos figuras geométricas planas, de tal forma que a cada uno de los puntos y rectas de una de ellas, le corresponden, respectivamente, un punto y una recta de la segunda figura.

Existe una relación entre los puntos de dos imágenes de la misma escena pero con diferentes puntos de vista, si dichos puntos caen en el mismo plano del mundo real (Figura 2.4)

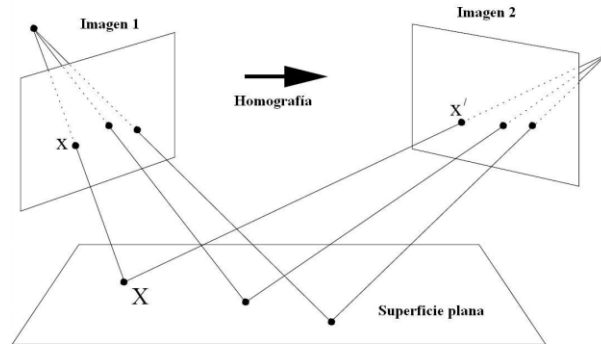


Figura 2.4: Relación entre dos imágenes de un mismo plano: Homografía.

La relación existente entre ambas imágenes viene determinada por una matriz de transformación 3x3, matriz Homografía H:

$$X' = HX \quad (2.6)$$

donde  $X$  es la proyección de un punto en una imagen y  $X'$  es la proyección del mismo punto en otra imagen situada en otro punto de vista. Tanto  $X$  como  $X'$  están expresadas en coordenadas homogéneas. En coordenadas homogéneas, todo punto bidimensional está definido por tres coordenadas. De tal modo que un punto de dimensiones  $x, y$ , como un punto de imagen, queda representado por la terna  $x/w, y/w, w$ . Se trata de ampliar el plano euclídeo bidimensional al plano proyectivo.

La expresión anterior (2.6) se desarrolla en:

$$\begin{bmatrix} x' \\ y' \\ w' \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} \begin{bmatrix} x \\ y \\ w \end{bmatrix} \quad (2.7)$$

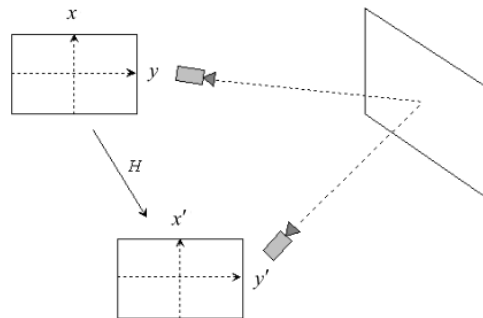


Figura 2.5: Explicación Homografía



De tal manera que, teniendo en cuenta las coordenadas homogéneas, los puntos de la nueva imagen son:

$$\begin{cases} x' = \frac{x'}{w'} = \frac{h_{11}x + h_{12}y + h_{13}}{h_{31}x + h_{32}y + h_{33}} \\ y' = \frac{y'}{w'} = \frac{h_{21}x + h_{22}y + h_{23}}{h_{31}x + h_{32}y + h_{33}} \end{cases}, \quad (2.8)$$

Como vemos, de la expresión que relaciona un punto de la imagen con el punto de la otra imagen a través de la homografía (2.6), resultan tres ecuaciones dependientes de los parámetros de la matriz de homografía H. Sin embargo, sólo dos de ellas (2.8) son linealmente independientes, ya que las coordenadas de un punto únicamente tienen dos grados de libertad.

Por este motivo, cada correspondencia de puntos determina dos restricciones sobre la matriz de homografía H.

Por otro lado, la matriz H tiene 9 parámetros (al tratarse de una matriz 3x3), pero queda definida salvo por un factor de escala ( $h_{33}$ ). De tal manera que presenta sólo 8 parámetros independientes. Por ello, es posible determinar la matriz de homografía H a partir de 4 puntos de correspondencia.

## 2.4 Casos de rectificación del plano de carretera:

La rectificación del plano es un problema condicionado por la cantidad y el tipo de información disponible (el número de puntos de fuga, sus relaciones, el conocimiento de la matriz de calibración, la posición relativa de la cámara con respecto al plano, la presencia de líneas para la determinación del punto de fuga, etc.). Por ello, se van a explicar a continuación diferentes métodos para el cómputo de la IPM [26]. Están ordenados en función del número de parámetros que se desconocen y se necesitan obtener para alcanzar la rectificación del plano de carretera.

### 2.4.1 Matriz de proyección conocida:

Se realiza la rectificación del plano de carretera a partir de las relaciones proyectivas entre las coordenadas de la imagen y el plano de carretera, asumiendo que se conocen todos los parámetros de la matriz de proyección: parámetros intrínsecos (matriz de calibración) y parámetros extrínsecos (rotación y traslación de la cámara con respecto al plano de carretera).

Considerando que los puntos del espacio 3D ( $X$ ) se encuentran en el plano de la carretera ( $Y = 0$ ), tenemos la siguiente relación proyectiva entre los puntos de la imagen ( $x$ ) y los del espacio 3D:

$$x = PX = \begin{pmatrix} p_1 & p_2 & p_3 & p_4 \end{pmatrix} \begin{pmatrix} X \\ 0 \\ Z \\ 1 \end{pmatrix} = \begin{pmatrix} p_1 & p_3 & p_4 \end{pmatrix} \begin{pmatrix} X \\ Z \\ 1 \end{pmatrix} = Hx' \quad (2.9)$$

Aplicando la homografía inversa según (2.10) se obtiene los valores de la imagen rectificadas:

$$x' = H^{-1}x \quad (2.10)$$

La generación de la imagen rectificadas transforma el punto de fuga en un punto en el infinito, que obviamente no puede ser representado. Por eso, es necesario seleccionar la región de la imagen original que va a ser transformada, con el objetivo de obtener una imagen rectificadas finita. De esta manera elegimos la profundidad de la carretera que vamos a mostrar en la imagen rectificadas.

## 2.4.2 Parámetros extrínsecos desconocidos:

En este segundo caso se asume el conocimiento de la matriz de calibración, pero el desconocimiento de los parámetros extrínsecos, es decir, no conocemos la posición relativa de la cámara con respecto al plano de la carretera. Este escenario es mucho más realista que el anterior, puesto que la rotación de la cámara varía.

La matriz de calibración se puede considerar estático en muchos casos. Se procede obteniendo una calibración inicial y asumiendo que permanecerá fija, y sólo computamos la rotación relativa de la cámara con respecto al plano de la carretera en cada instante. Además, la traslación de la cámara también puede ser considerada fija, puesto que la cámara se mueve conjuntamente con el vehículo.

De este modo, nuestro objetivo es obtener los valores de los ángulos de rotación: cabeceo (*pitch*,  $\theta$ ), guiñada (*yaw*,  $\gamma$ ) y alabeo (*roll*,  $\beta$ ).

### 2.4.2.1 Ángulo de alabeo desconocido:

Se simplifica el problema asumiendo que el ángulo de alabeo es nulo. Esta asunción se mantiene si la cámara está perfectamente instalada sin rotación con respecto al eje-Z.

De este modo, es suficiente con determinar el punto de fuga asociado con las líneas de carretera y determinar los ángulos  $\theta$  y  $\gamma$  según la fórmula (2.11):

$$\begin{aligned} \theta &= \arctan(v'_{z,2}) \\ \gamma &= \arctan\left(-\frac{v'_{z,1}}{\cos \theta}\right) \end{aligned} \quad (2.11)$$

Con este método obtenemos IPMs con líneas paralelas y de ancho constante, incluso en situaciones difíciles como un cambio de carril (varía el ángulo de guiñada) o un cambio de pendiente (varía el ángulo de cabeceo) [6].

#### 2.4.2.2 Rotación desconocida:

Aun desconociendo los tres ángulos de rotación, es posible conseguir la rectificación de la imagen dos puntos de fuga en la escena: el punto de fuga dominante ( $v_z$ ), que en la mayoría de las situaciones es posible extraer, y un segundo punto de fuga en la dirección del eje-X ( $v_x$ ), que requiere de la existencia de líneas discontinuas (Figura 2.6).

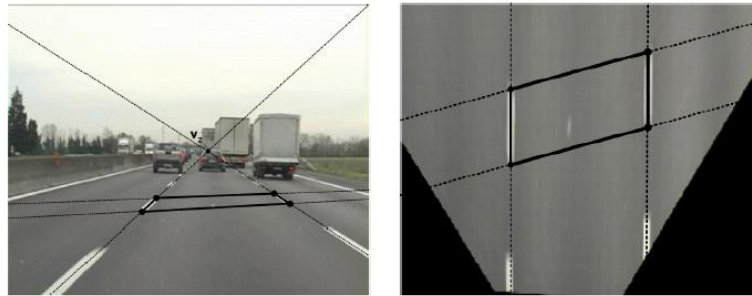


Figura 2.6: Detección segundo punto de fuga [26].

Una vez se han conseguido los dos puntos de fuga procedemos a obtener los ángulos de rotación que eran desconocidos.

Obtenemos el valor de cabeceo (*pitch*,  $\theta$ ) y guiñada (*yaw*,  $\gamma$ ) al igual que en el caso anterior, mediante la fórmula (2.11).

Y el ángulo  $\beta$  se obtiene a partir de la relación existente entre los dos puntos de fuga y la línea del horizonte [26].

Tras ello, ya tenemos todos los ángulos de rotación necesarios para construir la matriz de rotación y, con ella, generar la matriz de proyección para la rectificación de la imagen.

#### 2.4.3 Parámetros extrínsecos e intrínsecos desconocidos:

Esta última alternativa analiza la situación en la cual no se conocen ninguno de los parámetros que gobiernan el proceso de proyección.

##### 2.4.3.1 Sin calcular la matriz de proyección:

En este caso, además de no conocer los ángulos de rotación, ni disponemos ni somos capaces de computar la matriz de calibración de la cámara. A pesar de ello, somos capaces de realizar un rectificado de la imagen, aunque sin recuperar la información angular de la imagen, por lo que presentará una distorsión afín desconocida.

El método simplemente requiere la obtención del punto de fuga predominante de la imagen,  $v_z$ . Seguidamente, se seleccionan cuatro puntos de la imagen asumiendo que el ángulo de guiñada es nulo (esté método tendrá problemas en los movimientos de carril, en los que dicho ángulo no es nulo), tal y como se indica en la Figura 2.7. La selección de estos puntos se realiza uniando el punto de fuga con las esquinas inferiores de la imagen y seleccionando la intersección de esas líneas con dos líneas horizontales totalmente arbitrarias a diferentes alturas.

Mediante un conjunto de cuatro puntos de correspondencias, estimamos la homografía que representa la transformación entre los dos planos de imagen [11].

El resultado es una imagen rectificadora con una pequeña distorsión afín, que es prácticamente inapreciable visualmente.

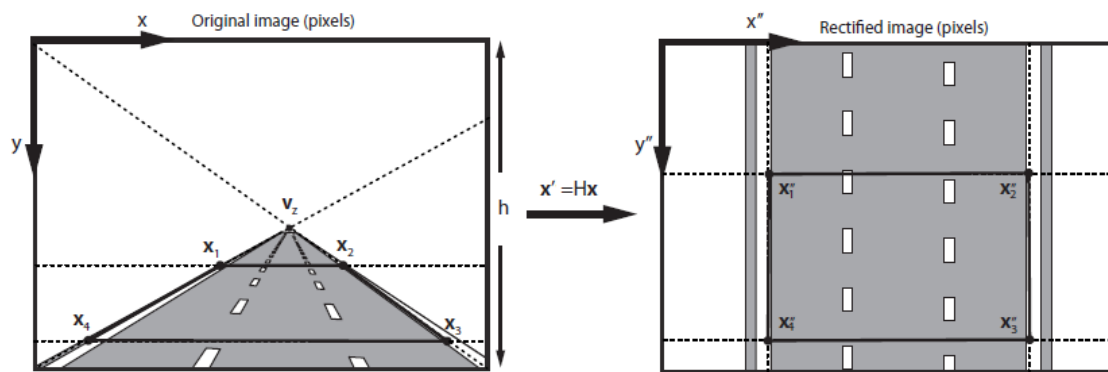


Figura 2.7: Computación de la homografía entre imágenes mediante correspondencia de cuatro puntos [26].

## 3. Diseño:

### 3.1 Estrategia general:

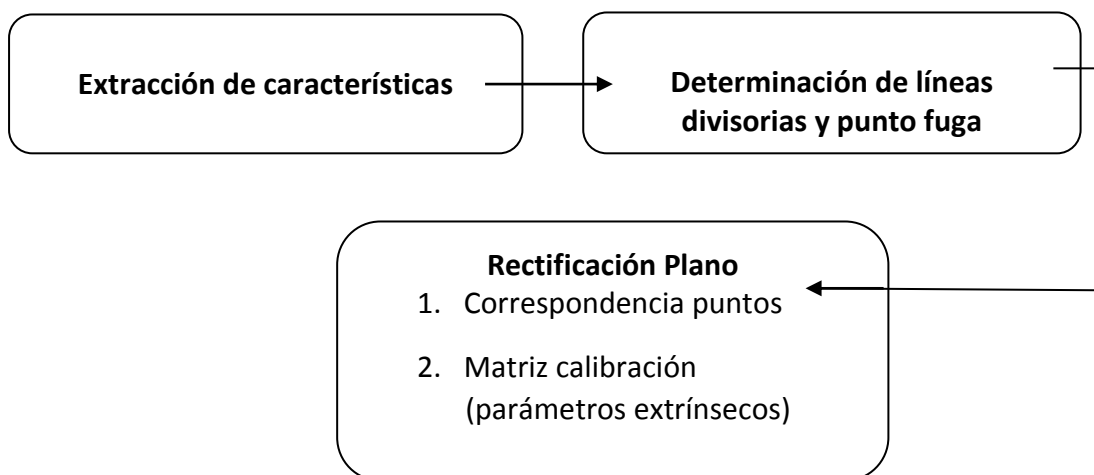
El objetivo final de nuestro sistema es conseguir una vista cenital de una imagen de carretera, a partir de una secuencia de imágenes grabadas con una cámara situada en la parte central de la luna del vehículo. Un ejemplo de esta vista lo podemos ver en la Figura 3.1.



Figura 3.1: Generación IPM: (a) Imagen original; (b) IPM (*Inverse Perspective Mapping*)

En esta vista cenital se elimina la distorsión de la perspectiva, consiguiendo numerosos beneficios: las líneas son paralelas en este dominio, tienen su anchura constante y real (salvo por un factor escala), las velocidades y posiciones de los vehículos se muestran sin distorsión, facilitando así todas las tareas de detección y localización de vehículos.

La obtención de forma automática, robusta y adaptativa de la vista cenital, vista de pájaro (*bird-view*) o IPM (*Inverse Perspective Mapping*) se lleva a cabo mediante el siguiente diagrama de bloques:



El primer bloque, Extracción de características, toma como entrada cada imagen de la secuencia de vídeo,  $I_n$ , y realiza la extracción de las características que nos interesan: las líneas divisorias de carril (*lane markings*) (a partir de ahora, por simplificar, las denominaremos líneas de carretera). Se proponen dos alternativas: aplicar un detector de bordes (Canny) o realizar un filtrado previo y una binarización. En ambas estrategias el objetivo es resaltar las líneas de la carretera de la parte inferior de la imagen, obteniendo una imagen de valor 1 en aquellos píxeles que pertenecen a línea de carretera y 0 en el resto. Con este primer paso conseguimos una imagen adaptada que sirva de entrada a nuestro detector de rectas, con la información únicamente necesaria para ello, evitando así la detección de posibles rectas no correspondientes a líneas de carretera.

El segundo bloque, Determinación de líneas divisorias y punto de fuga, toma como entrada la imagen binaria del apartado anterior y, mediante la aplicación de la Transformada de Hough, detecta las rectas presentes en la imagen, empleadas para la estimación del punto de fuga. La transformada de Hough, como explicaremos en detalle más adelante, se basa en un sistema de votación, obteniendo por tanto todas las posibles líneas contribuidas por todos los píxeles de la imagen. Los máximos de la transformada de Hough nos darán la información de las rectas más probables, que se corresponderán con las rectas correspondientes a las líneas de carretera. La determinación del punto de fuga se realizará a partir de la información de dichas rectas, siendo el punto en el que fugan las rectas paralelas presentes en la imagen. Se trata de un paso clave de nuestro sistema, puesto que una obtención robusta y precisa del punto de fuga afectará a la calidad de nuestra IPM. Por ello, también se realiza un seguimiento y filtrado del mismo, para garantizar una mayor suavidad y continuidad al sistema.

Por último, el tercer bloque se encarga de la rectificación del plano para conseguir la vista de pájaro o vista cenital, que es el objetivo final de nuestro sistema. La consecución de dicho objetivo se realiza mediante dos implementaciones distintas:

- La primera estima la imagen transformada a partir de una correspondencia de cuatro puntos [11]. Toma cuatro puntos de la carretera pertenecientes a dos rectas paralelas, formando un trapecio, y los sitúa formando un cuadrado en la IPM. De esta manera, se estima la matriz de transformación  $H$ , homografía, mediante el algoritmo DLT, que es la que nos proporciona la relación entre dos planos.

- Por otro lado, la segunda implementación tiene en cuenta los parámetros de calibración de la cámara para la transformación perspectiva [6]. Para ello, es necesario conocer información adicional: la matriz de calibración de la cámara (parámetros intrínsecos) y la posición relativa de la cámara (vector de traslación y matriz de rotación, parámetros extrínsecos). En esta segunda implementación aprovechamos la información del punto de fuga para realizar una estabilización de la IPM.

## 3.2 Extracción de características:

Tal y como se ha visto en la visión global del sistema, es necesario conocer y extraer cierta información de las imágenes de carretera para poder llevar a cabo la implementación de rectificación del plano. Esa información son las líneas divisorias (*lane markings*). Dado que estamos trabajando en un entorno muy concreto (con imágenes de carretera) podemos tener en cuenta varias asunciones que facilitarán en gran medida el proceso.

En el entorno en el que estamos trabajando y para el fin que buscamos es suficiente trabajar con imágenes en escala de grises. Las características que estamos buscando en las imágenes son líneas de carretera. Se conoce que esas líneas de carretera estarán presentes en gran medida en la parte inferior de la imagen. Además, en esa zona de la imagen las líneas de carretera se pueden aproximar como líneas rectas, incluso en presencia de cierta curvatura. También se conoce que el valor de los píxeles de las líneas de carretera será elevado, puesto que son de color blanco.

Para la consecución de este primer paso del sistema se proponen dos alternativas:

### 3.2.1 Aplicación detector de bordes:

La primera alternativa consiste en aplicar sobre la imagen cualquiera de los detectores de bordes. Se tratan de operadores que se basan en máscaras que aproximan, en mejor o peor medida, el gradiente de una imagen. Entre esos operadores se encuentran: Roberts, Prewitt, Sobel y Canny.

El mejor de ellos es el Detector de Canny, capaz de localizar los bordes de una imagen con alta precisión a partir del gradiente. Realiza un filtrado gaussiano previo (para eliminar ruido, puesto que el gradiente es muy sensible al ruido) y un adelgazamiento de contornos mediante una supresión de no máximos. En la Figura 4.2 se muestra el resultado de la aplicación de dicho detector sobre nuestra imagen de entrada.

Podemos observar que el resultado obtenido no es el más apropiado en nuestro caso, puesto que no extraemos las líneas de carretera en sí, sino los bordes de las mismas. La gran anchura de las líneas en la parte inferior de la imagen puede provocar que las rectas detectadas por la Transformada de Hough en el siguiente paso no se adapten de la mejor manera posible a las propias líneas de la carretera. En capítulos posteriores analizaremos el impacto negativo de emplear esta alternativa frente a la técnica del filtrado adaptado.



Figura 3.2: Detección de bordes con el operador Canny.

### 3.2.2 Filtrado adaptado:

Esta segunda alternativa trata de buscar una solución al problema adaptada a la situación en la que nos encontramos (extracción de líneas de carretera). Consta de los siguientes pasos:

#### A) Conversión de la imagen:

Convierte la imagen de entrada en una imagen en escala de grises. El motivo de esta conversión es por simplicidad en los algoritmos y puesto que la información relevante a las líneas de carretera se encuentra en la escala de grises.

#### B) Filtrado:

Asumiendo que las líneas de carretera son pulsos de alta intensidad rodeados de zonas oscuras, el análisis se realiza filtrando independientemente cada fila de la imagen. Con este filtrado conseguimos realzar los píxeles correspondientes a líneas de carretera [15].

$$y_i = 2x_i - (x_{i-\tau} + x_{i+\tau}) - |x_{i-\tau} - x_{i+\tau}|, \quad (3.1)$$

donde  $\tau$  es el parámetro del ancho que gobierna el proceso del filtrado.

Este filtrado produce respuestas altas en las posiciones de los píxeles que presentan un valor más alto que sus píxeles vecinos a la izquierda y a la derecha a una distancia  $\tau$ . El último término en (4.1) penaliza los casos en los que la diferencia entre los vecinos de la izquierda y la derecha sea alta. Por lo tanto, se obtienen respuestas altas en posiciones con valor elevado y cuyos vecinos presentan valores bajos y similares. La inclusión del último término hace a este filtro menos propenso a errores que otros detectores de líneas de carretera presentados en la literatura [6,28].

$\tau$  es un parámetro de diseño. La elección de su valor genera diferentes resultados en el filtrado de las líneas de carretera. Podemos ver el efecto para distintos valores de  $\tau$  en la Figura 3.3. Valores demasiado pequeños producen la detección de líneas más finas de lo que realmente son (Figura 3.3 (c)). Mientras que valores demasiado grandes producen detección de otros elementos no pertenecientes a las líneas de carretera, es decir, introducimos ruido en la detección (Figura 3.3 (d)). Un valor adecuado para nuestro entorno, obtenido empíricamente, es  $\tau=10$  (Figura 3.3 (b)).



En la Figura 3.4 vemos el efecto de dicho filtrado para una fila concreta de la imagen. En color azul tenemos los valores de la fila original, previa al filtrado, y en color rojo los valores de los píxeles tras el filtrado de dicha fila. Vemos cómo se acentúa la diferencia entre los píxeles de línea de carretera y el resto, consiguiendo así el objetivo que buscábamos: extraer (resaltar) las líneas de carretera de la imagen.

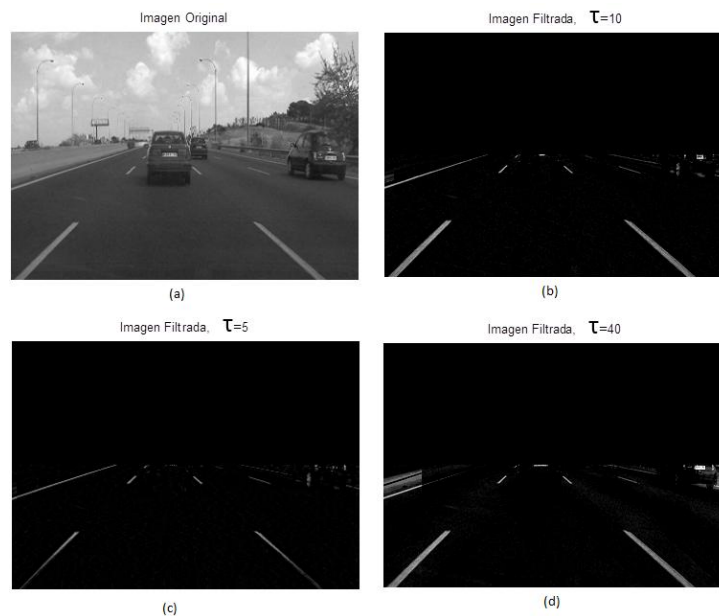


Figura 3.3: Filtrado líneas carretera: (a) Imagen original; (b) Imagen filtrada,  $\tau=10$ ; (c) Imagen filtrada,  $\tau=5$ ; (d) Imagen filtrada,  $\tau=40$

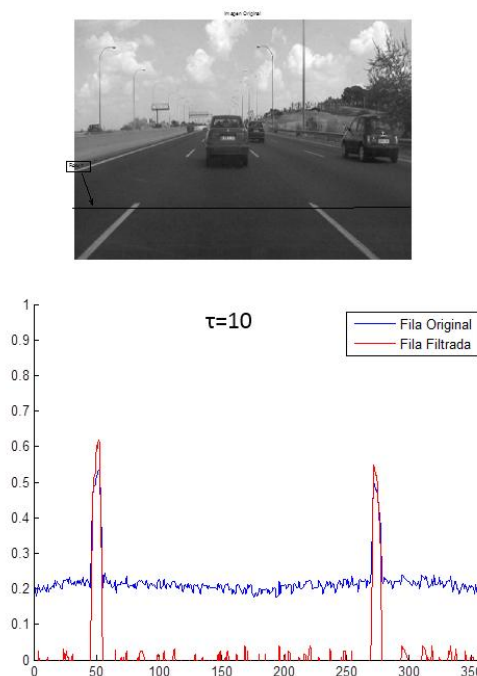


Figura 3.4: Efecto del filtrado para una fila.

### C) Binarización:

A pesar de no ser un paso estrictamente necesario para poder continuar con el sistema, se decide aplicar un proceso de binarización con el que eliminamos ruido y resaltamos aún más la diferencia entre los píxeles pertenecientes a las líneas de carretera y los píxeles restantes.

La binarización consiste en comparar cada píxel de la imagen con un determinado umbral (valor límite que determina si un píxel será de color negro o blanco). Los valores de la imagen superiores al umbral tomarán un valor 1 (blanco), mientras que el resto tomará valor 0 (negro).

Se emplea la umbralización como técnica básica de segmentación de imágenes bimodales. En la Figura 3.5 vemos el histograma bimodal de nuestra imagen de carretera.

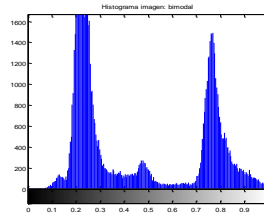


Figura 3.5: Histograma imagen bimodal.

El paso clave en la binarización es la selección del umbral global. Para ello empleamos la técnica de *Búsqueda iterativa del umbral que equidista de las medias superior e inferior* [27], que se desenvuelve en los siguientes pasos:

1. Selección del umbral inicial,  $T_{ini} = \frac{r_{max} - r_{min}}{2}$ .
2. Umbralización con  $T_{prev} = T_{ini}$  y obtención de los niveles medios  $\mu_O(T_{prev}), \mu_B(T_{prev})$  de los dos grupos de píxeles que separa el umbral (líneas de carretera y fondo).
3. Cálculo de un nuevo umbral,  $T_{new} = \frac{\mu_O + \mu_B}{2}$ .
4. Repetición de los pasos '2' y '3' hasta que la diferencia entre el nuevo umbral y el anterior sea menor que la unidad.

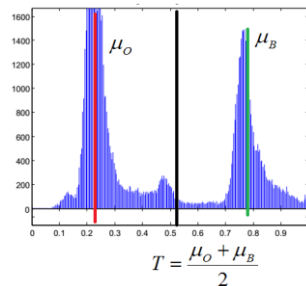


Figura 3.6: Búsqueda iterativa del umbral que equidista de las medias superior e inferior.

### 3.3 Determinación de líneas divisorias y punto de fuga:

El objetivo de este módulo es determinar el punto de fuga de la imagen en la dirección del desplazamiento y localizar las rectas que se corresponden con las líneas delimitadoras de los carriles de la carretera. Tanto para hallar el punto de fuga como para localizar las rectas aplicamos la **Transformada de Hough (HT)**.

La Transformada de Hough se encargará de obtener los parámetros (pendiente y ordenada en el origen) que modelan las rectas que se aproximan a las líneas de carretera. Así mismo, el método elegido para la determinación del punto de fuga es el ajuste de los máximos de la Transformada de Hough a una senoide.

#### 3.3.1 Transformada de Hough:

La Transformada de Hough, propuesta por Paul Hough en 1962, es una herramienta fundamental en las técnicas de visión artificial para la detección de líneas rectas, círculos y otras formas. Es un algoritmo basado en un esquema de votaciones que permite localizar formas regulares básicas en imágenes binarias aunque no estén perfectamente definidas. No se trata de una transformada discreta lineal, puesto que no es global, no es invertible y no respeta tamaños ni formas de las señales originales.

Consideramos un punto  $(x_i; y_i)$  y una recta definida mediante ese punto y su pendiente de la forma  $y_i = mx_i + b$ . Existen infinitas rectas que pasan por ese punto  $(x_i; y_i)$  y satisfacen la ecuación para infinitos valores de  $m$  y  $b$ . Si ahora consideramos el plano  $mb$  (llamado espacio de parámetros o espacio transformado) y escribimos esa ecuación de la forma  $b = -x_i m + y_i$ , tenemos la ecuación de una única recta para un punto  $(x_i; y_i)$  fijo. Además, para un segundo punto  $(x_j; y_j)$  también existe una única recta en el espacio de parámetros  $mb$  asociado a dicho punto. Esta segunda recta se cruza con la recta asociada a  $(x_i; y_i)$  en  $(m_i, b_i)$ , donde  $m_i$  representa la pendiente y  $b_i$  la ordenada en el origen de la recta que contiene a los puntos  $(x_i; y_i)$  y  $(x_j; y_j)$  en el plano original  $xy$ . Cada punto del espacio  $xy$  de la imagen se corresponde con una recta en el espacio de parámetros  $mb$ , así como cada punto en el espacio  $mb$  se corresponde con una recta en el espacio  $xy$ .

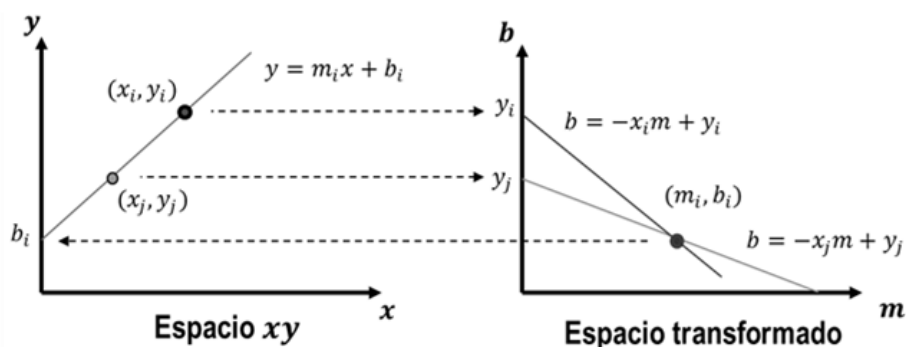


Figura 3.7: Correspondencia Espacio  $xy$  – Espacio  $mb$  transformado Hough [27].

El atractivo de la Transformada de Hough es que nos permite subdividir el espacio transformado en celdas acumuladoras, donde  $(m_{\min}, m_{\max})$  y  $(b_{\min}, b_{\max})$  son los rangos esperados para la pendiente y la ordenada en el origen. La celda de coordenadas  $(i, j)$  con un valor de acumulador  $A(i, j)$  se corresponde al cuadrado asociado con las coordenadas  $(m_i, b_j)$  del espacio transformado (ver imagen inferior). Para el cálculo de la HT (Transformada de Hough) se procede de la siguiente forma:

- Inicialmente se ponen todos los acumuladores a cero.
- Para cada punto  $(x_k, y_k)$  de la imagen, permitimos que  $m$  pueda tomar cualquier valor de los permitidos y calculamos  $b$  a partir de la ecuación  $b = -x_i m + y_i$ . Los valores resultantes del parámetro  $b$  se redondean hasta los permitidos. En ese caso, se incrementan en 1 las celdas correspondientes. Si para un valor de  $m_p$  resultó un valor  $b_q$ , se tiene que  $A(p, q) = A(p, q) + 1$ .
- Finalmente, un valor  $M$  en el acumulador  $A(i, j)$  representa que  $M$  puntos del plano  $xy$  caen sobre la recta  $y = m_i x + b_j$ . La precisión en la colinealidad de estos puntos depende del número de celdas del espacio de parámetros.

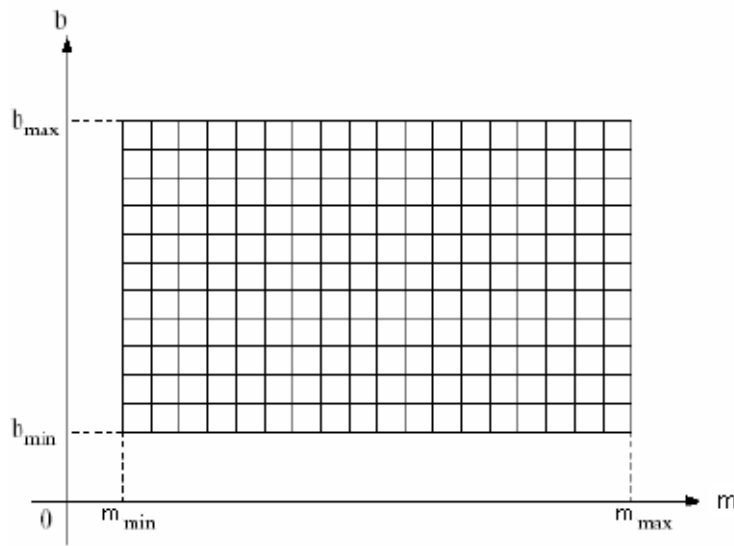


Figura 3.8: Acumulador plano  $mb$  de Hough

No obstante, la representación cartesiana de la recta presenta un problema: tanto la pendiente como la ordenada en el origen tienden a infinito conforme la recta se acerca a posiciones verticales. Para solucionar este problema, se propone emplear una representación alternativa: la representación polar (normal) de la recta:

$$r_i = y_i \sin \theta + x_i \cos \theta, \quad (3.2)$$

siendo  $r$ , la longitud de una normal desde el origen de la imagen (esquina superior izquierda) hasta la recta y  $\theta$ , el ángulo de  $r$  con respecto al eje X, los nuevos parámetros del espacio transformado, estando  $\theta$  acotado entre  $[-90^\circ, +90^\circ]$ .

En este nuevo espacio de parámetros, cada punto del plano  $xy$  se corresponde con una senoide en el plano  $(r, \theta)$ . Así mismo, si las curvas correspondientes a dos puntos en  $xy$  intersecan en un punto, la localización del mismo en el espacio de Hough se corresponde con una línea recta en el plano de la imagen que pasa por ambos puntos. En general, un conjunto de puntos que forman una línea recta en la imagen, producirán sinusoides que intersecarán en los parámetros  $(r, \theta)$  correspondientes a esa línea. Por ello, un problema de detección de puntos colineales se convierte en un problema de búsqueda de curvas concurrentes.

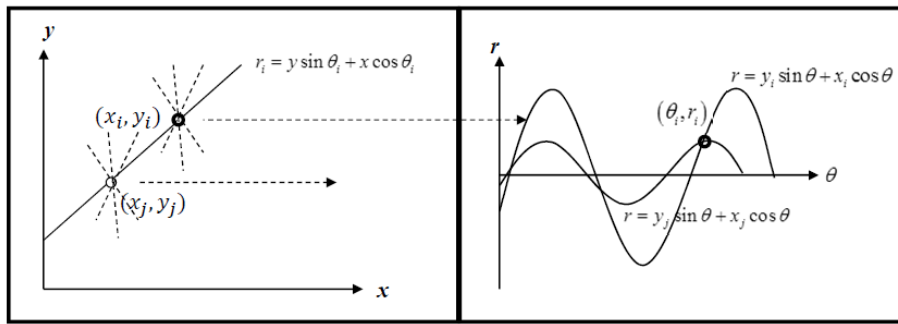


Figura 3.9: Correspondencia Espacio  $xy$  – Espacio  $mb$  transformado Hough [27].

El rango de representación de  $\theta$  podría ser teóricamente  $\pm 90^\circ$ . Sin embargo, dado que conocemos la inclinación aproximada que van a presentar las rectas que queremos detectar en la imagen, acotamos ese intervalo para así evitar detectar la línea del horizonte y posibles valores atípicos que introduzcan error en el cálculo del punto de fuga. Por lo tanto, el rango de  $\theta$  es  $\pm 80^\circ$ .

Por otro lado, el rango para  $r$  es  $\pm \sqrt{2D}$ , donde  $D$  es la diagonal de la imagen.

Podemos ver el resultado de aplicar la Transformada de Hough a la Figura 3.1 (a) en la Figura 3.10 (a).

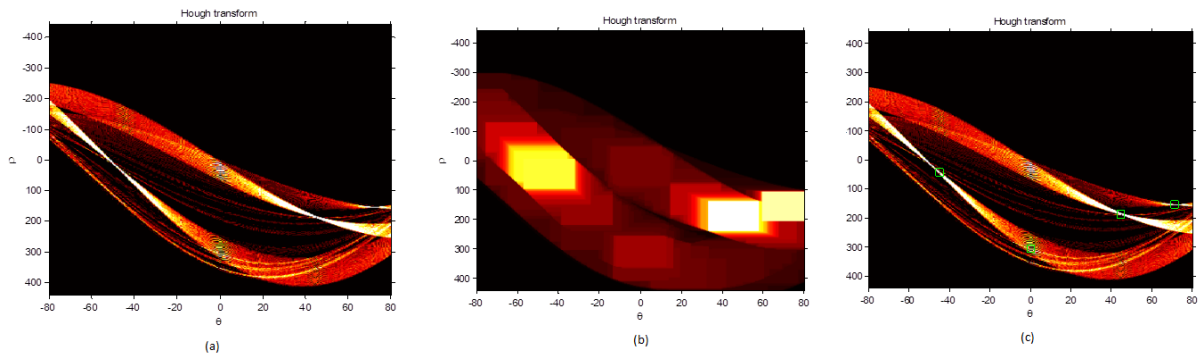


Figura 3.10: (a) Transformada de Hough; (b) Dilatación Transformada de Hough; (c) Detección de máximos en la Transformada de Hough.

### 3.3.2 Cálculo de máximos locales:

Tras aplicar la Transformada de Hough a la imagen tendremos que hallar los máximos de la misma, puesto que, como hemos explicado anteriormente, la Transformada de Hough se basa en un sistema de votaciones y será en los máximos de la Transformada dónde se encuentren los parámetros  $(r, \theta)$  de la recta más probable.

Para ello, es necesario aplicar un algoritmo de cálculo de máximos locales. En nuestro caso empleamos los operadores morfológicos para la obtención de máximos locales. Hacemos uso del operador morfológico de la dilatación. Dicho operador provoca una dilatación (ampliación) de los máximos de la imagen (Figura 3.11) con una vecindad determinada por el elemento estructurante. En este caso se elige un elemento estructurante rectangular, eliminando la detección de máximos cercanos en áreas rectangulares. Ello provoca el efecto que se muestra en la Figura 3.10 (b).

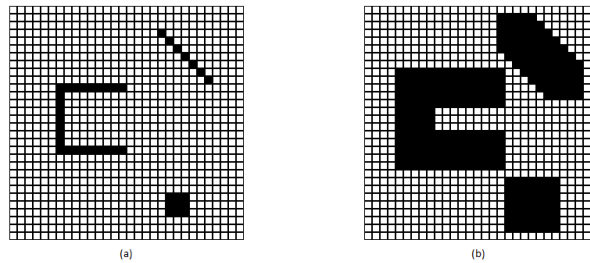


Figura 3.11: Aplicación Operador morfológico dilatación: (a) Imagen original; (b) Imagen dilatada

Lo único que no cambia tras la dilatación son los máximos (puesto que no se pueden dilatar más). Por ello, aplicando (3.3) conseguimos ceros en las posiciones de los máximos, es decir, conseguimos detectar los máximos de Hough.

$$I_{\max} = I - (I \oplus SE), \quad (3.3)$$

donde  $I$  es la imagen de entrada,  $SE$  el elemento estructurante con el que se dilata y  $I_{\max}$  la imagen con los máximos locales.

En la Figura 3.10 (c) observamos en verde los máximos detectados sobre la Transformada de Hough. El número de máximos detectados viene directamente determinado por el parámetro de entrada de nuestro sistema ( $num\_peaks$ ).

### 3.3.3 Ajuste de los máximos de la Transformada de Hough a una senoide:

Como dijimos, el objetivo de este bloque era la localización de las líneas de carretera y la determinación del punto de fuga. Una vez aplicada Hough y hallados sus máximos, tenemos los parámetros  $(r, \theta)$  de las rectas presentes en la imagen, consiguiendo localizar y detectar las rectas correspondientes a las líneas de carretera. Dichas rectas se muestran en verde en la Figura 3.12.

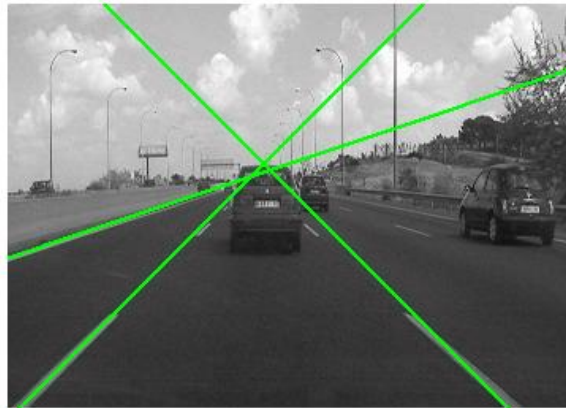


Figura 3.12: Resultado de la Transformada de Hough

El cálculo del punto de fuga, a pesar de no ser un paso necesario para la obtención de nuestro objetivo final (la vista cenital), se trata de uno de los puntos claves del sistema para conseguir resultados más robustos, fiables y estables.

El punto de fuga se define como el lugar geométrico en el cual las proyecciones de las rectas paralelas convergen. Dado que las líneas de la carretera son paralelas, dichas rectas convergerán en un punto en el plano de la imagen, siendo éste el punto de fuga (*vanishing point*).

Para su cálculo vamos a emplear el método de ajuste de los máximos de la Transformada de Hough a una senoide. Es un método que funciona muy bien en el caso de imágenes con un punto de fuga predominante, como es el caso de las imágenes de carretera.

La finalidad de este método es la obtención de un único punto de fuga (el punto de fuga predominante en la dirección de desplazamiento). Como sabemos, los máximos de la Transformada de Hough se corresponden con rectas de la imagen. Si varias rectas fugan en un mismo punto, en el espacio de Hough indica que varios máximos en el plano de Hough forman parte de una misma senoide. Por lo tanto, el procedimiento llevado a cabo es el de ajustar a una senoide los máximos detectados en Hough con los pasos anteriores. Esta senoide se corresponderá con el punto de fuga en el plano de la imagen.

Para encontrar esta senoide, realizamos el ajuste de los máximos de la Transformada minimizando el error cuadrático medio. En la Figura 3.13 (a) vemos la

sinusoide ajustada a los máximos de Hough y en la Figura 3.13 (b) la localización del punto de fuga correspondiente.

El método de ajuste de los máximos de la Transformada de Hough a una sinusoide proporciona resultados influenciados por los máximos que consideramos en Hough. Dado que la Transformada de Hough se basa en un sistema de votaciones (y debido a las oclusiones existentes en la imagen), es imposible garantizar que los primeros máximos identifiquen a las rectas correspondientes a las líneas de carretera. Por ello, el resultado va a depender del número de máximos (*num\_peaks*) que consideremos en dicho ajuste. Siempre que consideremos un número de máximos superior al número de líneas de la carretera presentes en la imagen, los resultados tendrán en cuenta rectas que no se corresponden con las líneas de carretera. Este hecho provoca que la sinusoide no se ajuste perfectamente a los máximos de la Transformada Figura 3.14 (a). La consecuencia de este mal ajuste es una mala determinación del punto de fuga (Figura 3.14 (b)).

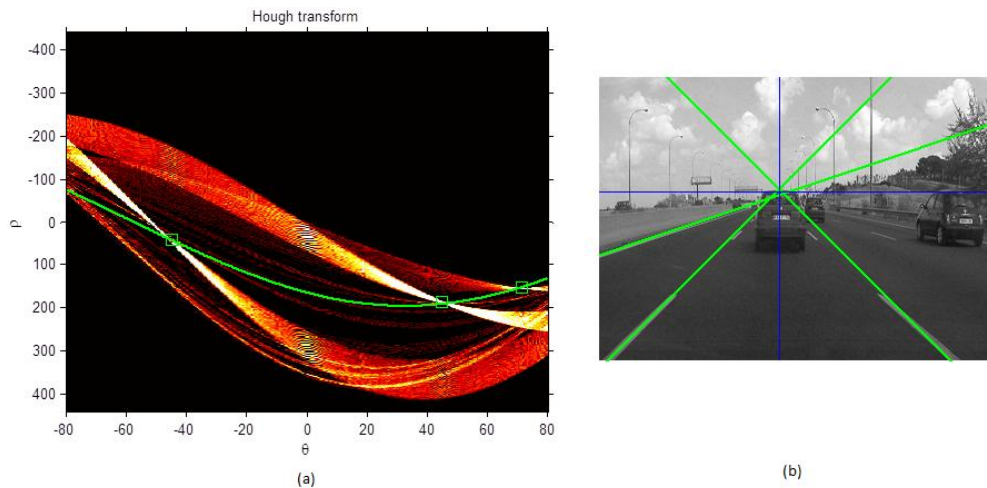


Figura 3.13: Ajuste a sinusoide: (a) Sinusoide en la Transformada de Hough; (b) Punto de fuga calculado a partir de la sinusoide.

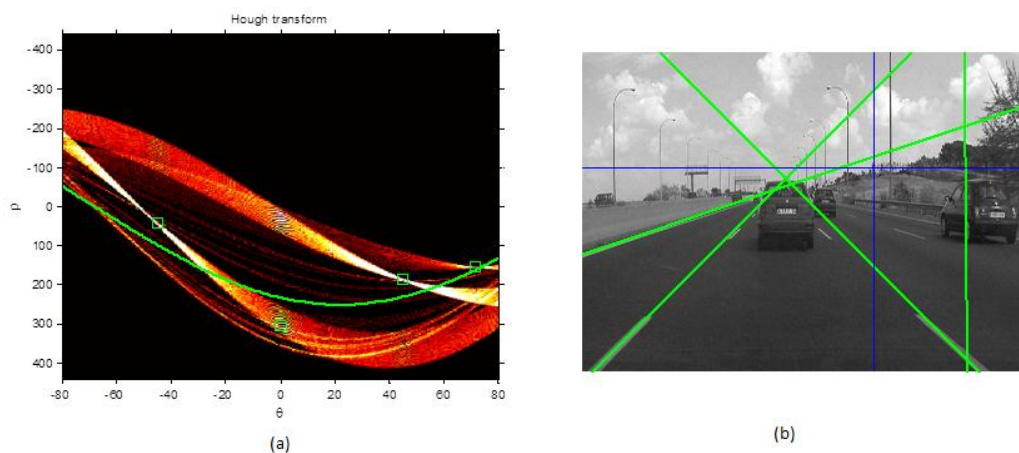


Figura 3.14: Mal ajuste a sinusoide: (a) Sinusoide en la Transformada de Hough; (b) Punto de fuga calculado a partir de la sinusoide.



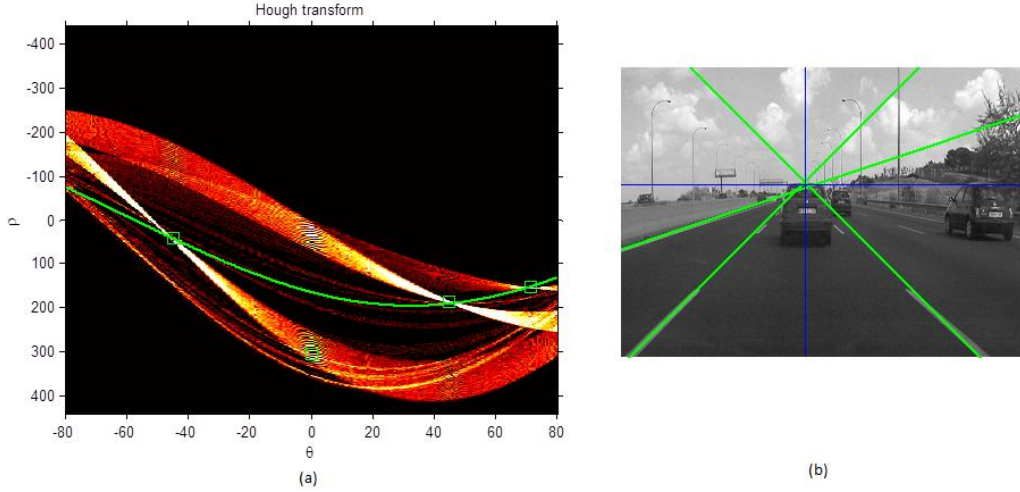


Figura 3.15: Ajuste a una senoide filtrado: (a) Senoide en la Transformada de Hough; (b) Punto de fuga calculado a partir de la senoide.

Para evitar este hecho, se propone emplear el conocimiento a priori que tenemos de la posición de las líneas de carretera en nuestra imagen. Sabiendo los rangos de inclinación que presentarán las líneas de la carretera presentes en la imagen, podemos acotar los datos obtenidos con la Transformada de Hough, dejando los máximos detectados que corresponden más probablemente a las líneas de carretera. De esta manera obtenemos un ajuste a una senoide adecuada para los máximos que han pasado el filtrado (Figura 3.15 (a)). Así mismo, el valor calculado del punto de fuga es el correcto (Figura 3.15 (b)).

### 3.3.4 Filtrado del punto de fuga:

El último paso realizado en este bloque es el filtrado del punto de fuga [6]. Se propone un filtrado paso-bajo con el objetivo de estabilizar el punto de fuga y eliminar las fluctuaciones presentes a lo largo del tiempo. A pesar de esperarse una variación prácticamente nula o de evolución suave en el punto de fuga, debido a errores en la extracción de características y detección de líneas de carretera, así como a vibraciones del sistema de adquisición, se observan saltos y fluctuaciones en su evolución (Figura 3.16 (a) y (c)). Para su suavizado se realiza el siguiente filtrado, considerando los  $m$  puntos de fuga previos para calcular la estimación del nuevo punto de fuga [6]:

$$\overline{v_n} = \overline{v_{n-1}} + \frac{1}{m}(v_n - v_{n-m}), \quad (3.4)$$

donde  $\overline{v_n}$  representa la estimación del punto de fuga, y  $v_n$  el punto de fuga calculado mediante el método descrito anteriormente.

Este filtrado temporal, junto con el hecho de emplear el conocimiento a priori de la inclinación de las rectas, asegura que los valores incorrectos del punto de fuga, debido a errores en el proceso de extracción y detección de líneas y a las propias

oclusiones, no afecten significativamente a la estimación final de  $\overline{v_n}$  y se consiga una variación suave y continua (sin saltos bruscos) del mismo.

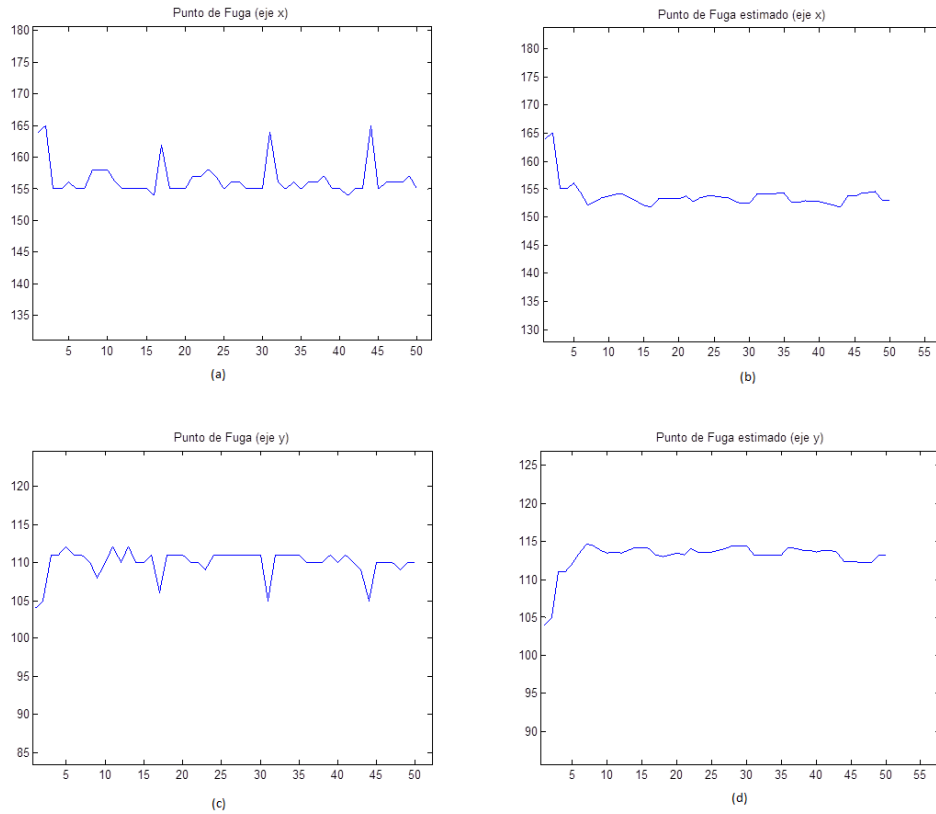


Figura 3.16: Variación punto de fuga y punto de fuga filtrado: (a) Variación eje-x punto de fuga; (b) Variación eje-x punto de fuga filtrado; (c) Variación eje-y punto de fuga; (d) Variación eje-y punto de fuga filtrado.

En la Figura 3.16 (b) y (d) se observa el suavizado de las fluctuaciones del punto de fuga a partir del instante  $m$  (siendo  $m=5$  en este caso). Conseguir este suavizado es un factor muy importante, puesto que la continuidad de las IPM generadas en el sistema dependerá de ello.

En la Figura 3.17 podemos ver el punto de fuga calculado (en línea continua,  $v_n$ ) y el punto de fuga filtrado (en línea discontinua,  $\overline{v_n}$ ).

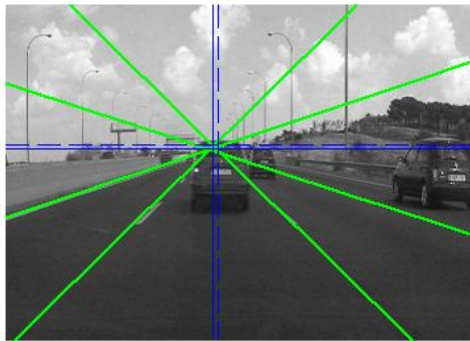


Figura 3.17: Representación del punto de fuga (línea continua) y punto de fuga filtrado (línea discontinua).

### 3.4 Rectificación del Plano:

La rectificación del plano es un problema condicionado a la cantidad y tipo de información disponible: el punto de fuga, conocimiento de la matriz de calibración, la posición relativa de la cámara con respecto al plano, oclusiones de las líneas de carretera, etc.

Por ese motivo, en este trabajo se han desarrollado 2 aproximaciones diferentes para rectificar planos en el escenario de la carretera, que requieren diferente tipo de información para su funcionamiento: **cómputo de la homografía mediante correspondencia entre puntos** y **cálculo de la IPM mediante la matriz de proyección**.

#### 3.4.1 Cómputo de la Homografía mediante correspondencia entre puntos:

El objetivo de este módulo es realizar una transformación de un plano de imagen en otro plano de imagen: transformar la imagen de carretera capturada con la cámara en la imagen que se observa desde otro punto de vista: imagen vista cenital, vista de pájaro (*bird-view*) o IPM (*Inverse Perspective Mapping*).

La transformación entre dos planos está gobernada por una Homografía. La forma elegida para obtener dicha matriz  $H$  es mediante la correspondencia de 4 puntos.

Dadas 4 correspondencias de puntos, denotadas en coordenadas homogéneas como  $(x, y, 1) \leftrightarrow (x', y', w')$  y en las cuales no hay 3 puntos colineales, llegamos a 8 ecuaciones que son suficientes para resolver los parámetros de  $H$ , salvo un factor de escala (Ecuación 2.7).

Los pasos llevados a cabo en esta implementación son:

##### A) Obtención de líneas de carretera de nuestro carril:

A la entrada de este bloque tenemos la imagen con las líneas de carretera detectadas y parametrizadas, y el punto de fuga localizado. No obstante, tal y como habíamos mencionado anteriormente, la Transformada de Hough detecta cualquier línea presente en la imagen, mientras que para nuestro caso únicamente nos interesan las líneas del carril en el que nos encontramos (carril propio, *own lane*), situadas una a cada lado del centro de la imagen.

Realizar el filtrado y selección de dichas rectas es una tarea sencilla, puesto que únicamente es necesario introducir información conocida: el ángulo de las líneas de nuestro carril será siempre el mismo dentro de un intervalo (recta izquierda:  $20^\circ$  a  $65^\circ$ ; recta derecha:  $-20^\circ$  a  $-65^\circ$ ). No obstante, no siempre es posible obtener dichas líneas con Hough, debido a las oclusiones producidas por la presencia de varios vehículos y la discontinuidad de las líneas. Este hecho hace que sea necesario llevar a

cabo una **gestión de oclusiones** para la correcta obtención de las líneas del carril en el que nos encontramos.

- La primera alternativa para esta gestión de oclusiones es utilizar información del instante anterior. Esta solución no parece muy adecuada en los casos en los que las oclusiones permanecen a lo largo de varias imágenes consecutivas y se produce un desplazamiento horizontal del vehículo, puesto que las líneas detectadas cada vez diferirán más de la realidad. Analizaremos en más detalle el impacto de esta solución en el próximo capítulo.
- Una alternativa avanzada es emplear más información presente en la imagen para inferir las líneas de interés: puesto que Hough detecta varias líneas presentes en la imagen, será capaz de detectar las líneas de los carriles adyacentes. Asumiendo que la anchura del carril permanece constante (hecho totalmente asumible salvo por pequeñas variaciones), podemos extrapolar la información a partir de otras líneas de la carretera. Con esta segunda variante obtenemos una mejor detección de las líneas de nuestro carril. Además, mejoramos los resultados incluyendo un filtrado paso-bajo (media) de la anchura del carril, con el que tenemos en cuenta las variaciones de la anchura del carril a lo largo del tiempo.

#### *B) Reestimación de rectas:*

Tras la detección de las líneas de nuestro carril empleamos la información del punto de fuga para realizar una reestimación de dichas rectas.

Dado que nuestro punto de fuga ha sido calculado empleando más de dos rectas detectadas y ha sido filtrado para obtener una mejor estimación del mismo, podemos considerar dicha información fiable. Por ello, parece lógico aprovechar esta información para mejorar la estimación de las rectas. Tras este paso, tenemos detectadas las rectas que se aproximan a las líneas de carretera de nuestro carril. En la Figura 3.18 se observa las líneas detectadas con Hough (verde) y las líneas reestimadas con el punto de fuga (rojo).



Figura 3.18: Reestimación rectas (líneas rojas) a partir del punto de fuga (azul)

### C) Selección de los puntos de correspondencias:

El siguiente paso del proceso es seleccionar 4 puntos arbitrarios de las dos líneas anteriores. Seleccionamos 4 puntos que sabemos que son paralelos 2 a 2, formando un trapecio (Ver Figura 3.19 (a)). Asumiendo que las líneas de carretera son perpendiculares a la horizontal de la imagen, elegimos dos alturas diferentes en la imagen, que determinarán dos rectas horizontales que cortarán a las líneas de carretera de nuestro carril, obteniendo los puntos A, B, C y D.

Los otros 4 puntos correspondientes, de la imagen con vista cenital, se seleccionarán también de forma arbitraria, con la restricción de que tienen que formar un rectángulo, puesto que dichos puntos tienen que estar alineados paralelamente (Ver Figura 3.19 (b)). La separación elegida en dichos puntos controla la distancia de observación de la imagen: si los seleccionamos muy juntos se mostrará prácticamente toda la imagen, dando sensación de una distancia de observación lejana. Mientras que si elegimos una gran separación, únicamente se mostrará una parte concreta de la imagen, dando una sensación de distancia de observación cercana. Dicho efecto se muestra en la Figura 3.20.

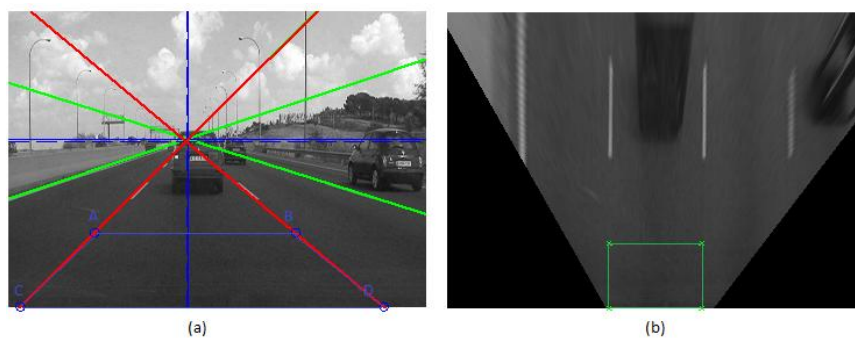


Figura 3.19: Selección de puntos de correspondencia: (a) Puntos formando un trapecio en la imagen original; (b) Puntos formando un cuadrado en la IPM

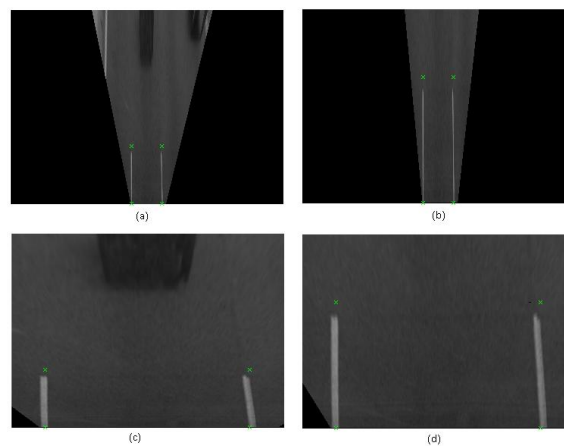


Figura 3.20: Selección de puntos de correspondencia en IPM: (a) Poca separación en ambos ejes; (b) Separación grande en el eje-y; (c) Separación grande en el eje-x; (d) Separación grande en ambos ejes.

*D) Obtención de la matriz homografía H (algoritmo DLT):*

Mediante el algoritmo DLT (*Direct Linear Transformation*) obtenemos la homografía  $H$ . El algoritmo DLT computa la ecuación  $X' = HX$  en términos del producto vectorial como  $X'_i \times (HX_i) = 0$ , siendo  $X'_i = (x'_i, y'_i, z'_i)$  y  $X_i = (x_i, y_i, z_i)$ .

La expresión  $X'_i \times (HX_i) = 0$  se transforma en  $Ah = 0$  (Ver Anexo A.1). Dicha ecuación se resuelve realizando la descomposición de valores singulares (SVD) de la matriz  $A$  :

$$A = \begin{bmatrix} 0 & 0 & 0 & -(z'_1x_1) & -(z'_1y_1) & -(z'_1z_1) & (y'_1x_1) & (y'_1y_1) & (y'_1z_1) \\ (z'_1x_1) & (z'_1y_1) & (z'_1z_1) & 0 & 0 & 0 & -(x'_1x_1) & -(x'_1y_1) & -(x'_1z_1) \\ 0 & 0 & 0 & -(z'_2x_2) & -(z'_2y_2) & -(z'_2z_2) & (y'_2x_2) & (y'_2y_2) & (y'_2z_2) \\ (z'_2x_2) & (z'_2y_2) & (z'_2z_2) & 0 & 0 & 0 & -(x'_2x_2) & -(x'_2y_2) & -(x'_2z_2) \\ 0 & 0 & 0 & -(z'_3x_3) & -(z'_3y_3) & -(z'_3z_3) & (y'_3x_3) & (y'_3y_3) & (y'_3z_3) \\ (z'_3x_3) & (z'_3y_3) & (z'_3z_3) & 0 & 0 & 0 & -(x'_3x_3) & -(x'_3y_3) & -(x'_3z_3) \\ 0 & 0 & 0 & -(z'_4x_4) & -(z'_4y_4) & -(z'_4z_4) & (y'_4x_4) & (y'_4y_4) & (y'_4z_4) \\ (z'_4x_4) & (z'_4y_4) & (z'_4z_4) & 0 & 0 & 0 & -(x'_4x_4) & -(x'_4y_4) & -(x'_4z_4) \end{bmatrix}$$

Una vez conocida la matriz  $A$  es posible conocer los parámetros de la matriz homografía  $H$ , puesto que precisamente el núcleo de  $A$  es la solución de  $h$ .

Por lo tanto, debemos realizar la SVD de la matriz  $A$  :

$$A = UDV^t,$$

donde la última columna de  $V$  es el núcleo de  $A$ , que se corresponde con el vector  $h$ . De esta manera, obtenemos la matriz homografía  $H$  :

$$H = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix}$$

*E) Aplicación de la Homografía:*

Una vez obtenida la matriz, podemos realizar el cambio de perspectiva de la imagen, obteniendo la vista cenital. El último paso del proceso es, por lo tanto, la aplicación de la homografía.

Se parte de dos imágenes, la imagen original cuyos píxeles tienen coordenadas  $x(x_i, y_i, 1)$  y la imagen transformada cuyos píxeles tienen coordenadas  $x'(x'_i, y'_i, 1)$ .

La aplicación directa de la transformación  $X' = HX$  puede dar lugar a coordenadas de los píxeles  $X'$  no enteras, siendo necesario interpolar, pero sin saber a priori el valor de los píxeles vecinos.

Por este motivo, la manera de proceder en estos casos es aplicar la transformación inversa (Figura 3.21). Se parte de las coordenadas de la imagen transformada a la cual se le multiplica por la matriz homografía inversa para obtener las coordenadas de la imagen original:

$$X = H^{-1} X'$$

De esta manera, aplicamos sobre la imagen 2 la homografía inversa ( $H^{-1}$ ), obteniendo las coordenadas de la imagen 1. En esta imagen sí son conocidos o fácilmente interpolables los valores de todos los píxeles, puesto que se trata de la imagen de partida.

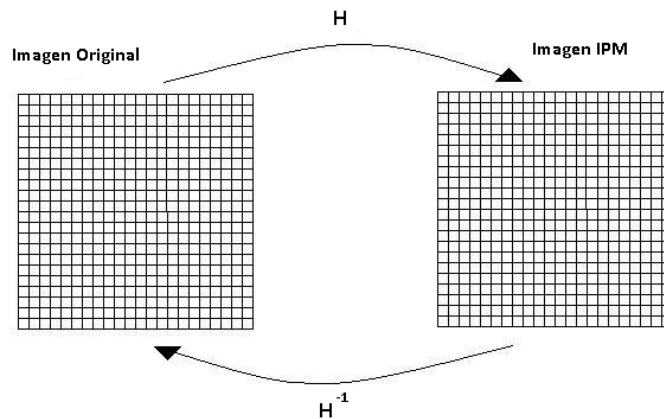


Figura 3.21: Homografía ( $H$ ) y Homografía inversa ( $H^{-1}$ )

Se procede de la siguiente manera: vamos recorriendo los píxeles de la imagen 2, aplicando la matriz  $H^{-1}$  y obteniendo las coordenadas en la imagen 1. A partir de esas coordenadas se busca el valor del píxel en la imagen 1 y se asigna al píxel de la imagen 2 del que se partió. También es posible que en la operación  $X = H^{-1} X$  se obtengan coordenadas de píxeles no enteras, teniéndose que interpolar, empleando cualquier técnica de interpolación, los valores de los píxeles a partir de los valores de la imagen original. Particularmente en este trabajo, se utilizan directamente el redondeo (redondeo a la coordenada de píxel entera más cercana) y la interpolación bilineal. Esta última es una extensión de la interpolación lineal para el caso de funciones de dos variables, como son en este caso las matrices de las imágenes (Figura 3.22).

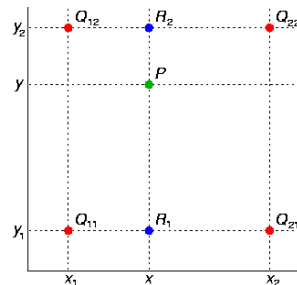


Figura 3.22: Interpolación bilineal

Se realiza primero una interpolación lineal en la dirección  $x$  ( $R_1$ ), seguido de una interpolación lineal en la dirección  $y$  ( $R_2$ ), generando finalmente el valor interpolado  $P$ .

Observamos el resultado final del proceso, la imagen rectificadora, en la Figura 3.23, tanto para el redondeo (Figura 3.23 (a)) como para la interpolación bilineal (Figura 3.23 (b)). Se aprecia claramente unos bordes mejor definidos en el caso de la interpolación bilineal.

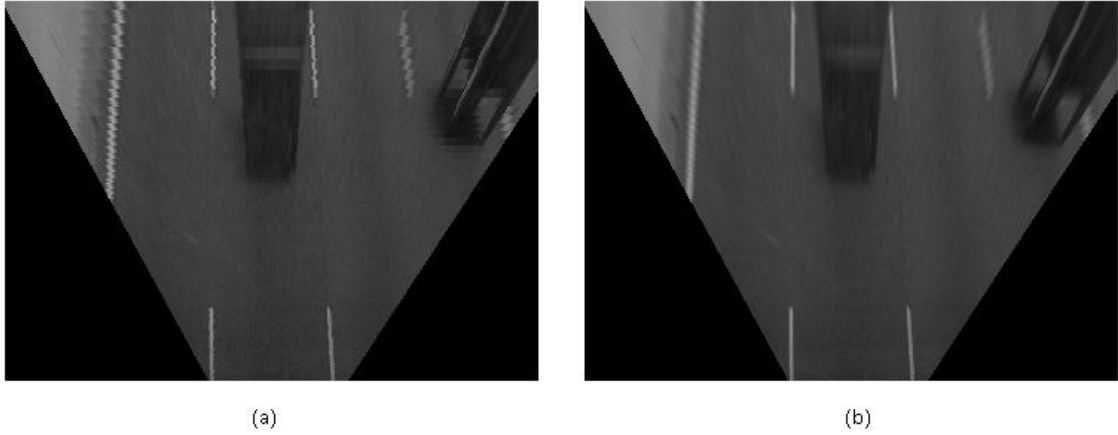


Figura 3.23: Imagen Rectificado mediante cuatro Puntos de Correspondencias: (a) Redondeo; (b) Interpolación bilineal

### 3.4.2 Cálculo de la IPM mediante la Matriz de Proyección:

Alternativamente al método de la Homografía por correspondencia de cuatro puntos explicado antes, se propone un segundo método para la consecución de la rectificación del plano.

Dicho método aplica una transformación de las coordenadas no-lineal, teniendo en cuenta las relaciones proyectivas entre las coordenadas del plano de carretera y el plano de la imagen, asumiendo el conocimiento de los parámetros de la matriz de proyección.

Con este método se consigue una IPM estabilizada y robusta, haciendo uso de la información del punto de fuga. Para realizar la rectificación del plano de carretera vamos a emplear el conocimiento de la matriz de proyección  $P$ .

Considerando que los puntos del espacio 3D se encuentran en el plano de la carretera (plano  $XZ$ , con coordenada  $Y = 0$ ) la relación proyectiva se simplifica a:

$$x = PX = \begin{pmatrix} p_1 & p_2 & p_3 & p_4 \end{pmatrix} \begin{pmatrix} X \\ 0 \\ Z \\ 1 \end{pmatrix} = \begin{pmatrix} p_1 & p_3 & p_4 \end{pmatrix} \begin{pmatrix} X \\ Z \\ 1 \end{pmatrix} \quad (3.5)$$



Los pasos llevados a cabo en este proceso son:

*A) Obtención de la matriz de proyección:*

Con el conocimiento de la matriz de proyección podremos ser capaces de realizar la rectificación del plano. En nuestro caso, asumimos el conocimiento de la matriz de calibración, es decir, se trata de un parámetro dado como entrada al sistema. Así mismo, se asume que los ángulos de rotación son nulos, es decir, se considera una instalación de la cámara dentro del vehículo perfecta. También consideramos la traslación del sistema de coordenadas de la cámara fijo con respecto al sistema de coordenadas del mundo.

*B) Utilización de la matriz de proyección: retro-proyección*

La IPM se construye haciendo uso de esa matriz y la máxima distancia de carretera a mostrar,  $Z_{\max}$ .

Este valor es transformado a coordenadas de la imagen, de tal manera que define la región de la imagen que va a ser retro-proyectada al mundo 3D (sólo la parte baja de la imagen es transformada).

Para la generación de la imagen IPM se realizan los siguientes pasos [6]:

1. Se calculan las 4 esquinas de la región de interés a partir de  $Z_{\max}$  y se retro-proyectan al mundo real en el plano  $Y = 0$ .
2. Se divide el plano de carretera en celdas de tal manera que cada una se proyecta en la imagen IPM en un único punto.
3. Se aplica la proyección a cada celda, obteniendo el píxel de la imagen original. El valor de dicho píxel rellena el correspondiente de la IPM. En el cálculo de la proyección es posible obtener valores no enteros de píxeles, procediéndose a interpolar su valor a partir de los adyacentes empleando los mismos métodos que en la estrategia anterior (redondeo e interpolación bilineal).

En la Figura 3.24 vemos la IPM obtenida. Se puede observar que los resultados no son los esperados: las rectas no son paralelas del todo. Esto se debe a que en la situación real, los ángulos  $\theta$  y  $\gamma$  nunca son nulos. Siempre existen pequeñas variaciones del ángulo  $\theta$  debido a pendientes o baches. Además, la instalación de la cámara (que no es perfecta) y el movimiento de cambio de carril influyen en la variación de dichos ángulos de rotación.

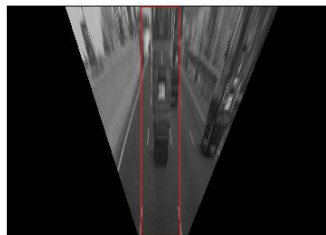


Figura 3.24: IPM sin estabilización

### C) Estabilización de la IPM

En este último paso se considera que los ángulos de rotación no son nulos, excepto el ángulo de alabeo ( $\beta$ ) que sí se considera nulo, puesto que esta asunción se mantiene si la cámara está instalada cuidadosamente sin rotación con respecto al eje-Z.

Por lo tanto, nuestro objetivo es obtener los valores de los ángulos de balanceo ( $\theta$ ), guiñada ( $\gamma$ ) para la construcción de la matriz  $R$ .

Podemos computar los ángulos  $\theta$  y  $\gamma$  únicamente a partir del punto de fuga  $v_z$ .

Partimos del punto de fuga calculado en el módulo anterior (en coordenadas de la imagen) y lo retro-proyectamos en el sistema de referencia de la cámara:

$$v'_z = K^{-1}v_z \quad (3.6)$$

Los ángulos  $\theta$  y  $\gamma$  se pueden calcular directamente como:

$$\begin{aligned} \theta &= \arctan(v'_{z,2}) \\ \gamma &= \arctan\left(-\frac{v'_{z,1}}{\cos \theta}\right) \end{aligned} \quad (3.7)$$

Estas expresiones provienen de considerar que el punto en el infinito se corresponde con el punto de fuga  $v_z$ , y su proyección en la imagen es:

$$v_z = K(R | -R \cdot t) \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} = K \begin{pmatrix} -\tan \gamma \cdot \cos \theta \\ \tan \theta \\ 1 \end{pmatrix} \quad (3.8)$$

Una vez obtenidos los valores del cabeceo (*pitch*) y guiñada (*yaw*) a partir de la estimación del punto de fuga (3.15), tenemos corregida la matriz de rotación y, con ella, la matriz de proyección. Se procede de igual manera que sin la estabilización, siguiendo los pasos de  $B$ .

Podemos ver la comparativa entre ambas en la Figura 3.25: teniendo la IPM estabilizada (a) las rectas paralelas, mientras que en la IPM sin estabilizar (b) sus líneas de carretera presentan cierta inclinación, no siendo paralelas del todo.

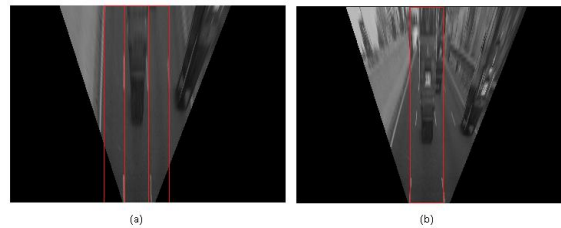


Figura 3.25: Comparativa IPM estabilizada (a) y IPM No estabilizada (b)

## 4. Resultados:

El objetivo de este apartado es mostrar una evaluación del rendimiento del sistema. Para ello, se van a mostrar los resultados obtenidos con diferentes variantes de los parámetros configurables. Los puntos a analizar son tanto la determinación de las líneas de carretera y obtención del punto de fuga, como la generación de la imagen rectificada (IPM).

En relación con el punto de fuga, se va a comparar los valores (posición de los píxeles) obtenidos con nuestro sistema frente a los valores teóricos (calculados manualmente que constituyen la referencia o *ground-truth*). También será objeto de estudio la variabilidad que presente dicho punto de fuga a lo largo del tiempo, dado que esas fluctuaciones tienen un impacto directo en la suavidad y continuidad de las IPMs generadas.

En cuanto a la IPM, al tratarse de secuencias de imágenes, la evaluación consiste en la comparativa entre la imagen generada y una imagen de referencia que consideramos perfecta o ideal (*ground-truth*). En dicha comparación se estudiará la inclinación de las líneas de carretera (comprobar que realmente son paralelas), anchura de los carriles (debe ser constante) y la profundidad de imagen mostrada.

Dado que es impracticable obtener una rectificación del plano de forma teórica, consideramos nuestra imagen de referencia (*ground-truth*) a la calculada a partir de la generación de la IPM conociendo la matriz de proyección y estimando los parámetros extrínsecos para su estabilización a partir de un punto de fuga determinado manualmente. Es decir, aplicamos el segundo método de rectificación de la imagen, pero introduciendo información teórica del punto de fuga.

Se ha llevado a cabo pruebas con dos tipos de secuencias:

- Secuencias sintéticas: generada mediante la matriz de proyección. Las pruebas realizadas con este tipo de secuencia sirven para validar ambos métodos y verificar la bondad de las imágenes de referencia (ver Figura 4.1).
- Secuencias reales: a partir de imágenes reales capturadas con la cámara y puntos de fuga anotados manualmente, obtenemos las imágenes de referencia.

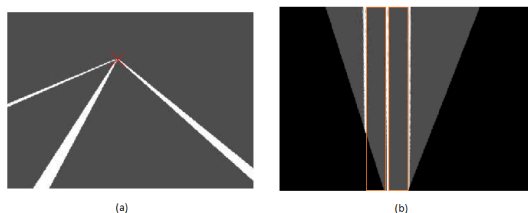


Figura 4.1: Verificación de la bondad del método 2 a partir de una secuencia sintética (a).

## 4.1 Impacto en el punto de fuga:

Comenzaremos analizando en detalle el impacto que tiene en el punto de fuga la determinación de algunos de los parámetros del sistema. Los casos de estudio están organizados según los módulos del sistema.

### 4.1.1 Parámetros de la extracción de características

#### a) Operador de Canny:

El primer parámetro del sistema nos permite elegir el operador de Canny para la detección de bordes de la imagen. Tal y como dijimos en el capítulo anterior, este operador extrae los bordes presentes en la imagen, sin centrarse únicamente en los píxeles que forman parte del interior de las líneas de carretera. Las rectas que detectamos vienen determinadas por todos los píxeles de borde de la imagen (parte inferior), con lo que son rectas que no se tienen por qué corresponder con las líneas divisorias de la carretera, sino también con los bordes de los vehículos presentes en la escena. Este hecho provoca una determinación inexacta del punto de fuga.

En la Figura 4.2 se observa una baja diferencia de valores (tanto para el eje-x como para el eje-y) entre el punto de fuga estimado y el punto de fuga teórico. Canny ofrece muy buenos resultados en situaciones en las que no hay otros elementos (como coches) en la escena. En los casos en los que aparecen otros elementos en la escena y sus bordes son detectados, los resultados calculados difieren en mayor medida de los teóricos (Figura 4.3). Podemos ver los valores concretos en las tablas A.2.1 y A.2.2 del Anexo A.2.

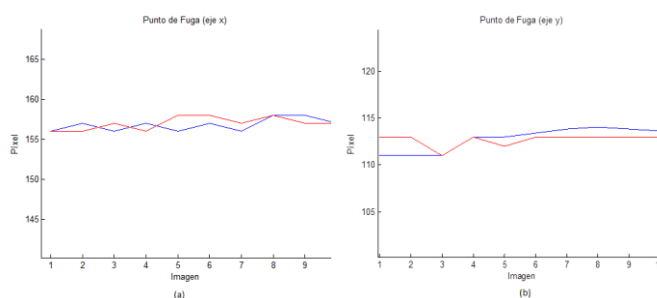


Figura 4.2: Comparación punto de fuga calculado (azul) y punto de fuga teórico (rojo) para el eje-x (a) y el eje-y (b), buena situación.

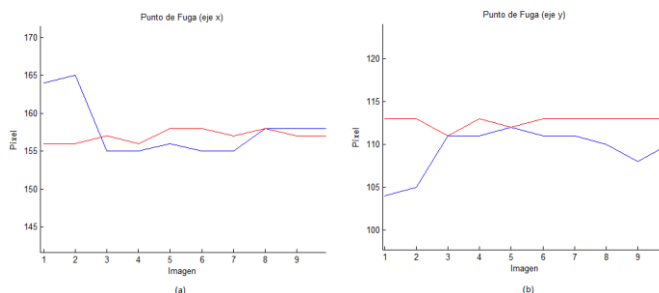


Figura 4.3: Comparación punto de fuga calculado (azul) y punto de fuga teórico (rojo) para el eje-x (a) y el eje-y (b), mala situación.

#### b) Filtrado adaptado:

Con este método obtenemos una detección de las líneas de carretera más aproximada, puesto que tiene en cuenta todos los píxeles que forman parte de la línea (tanto bordes como internos). En la Figura 4.4 se observan las líneas detectadas. Esta mejor aproximación tiene relación directa con una mejor determinación del punto de fuga (Figura 4.5). Las diferencias obtenidas en estos casos, a pesar de no ser tan buenas como en las mejores situaciones de Canny, son más estables ante casi cualquier situación (aparición de otros elementos en el plano de carretera). En la tabla A.2.3 tenemos los valores exactos de diferencia.

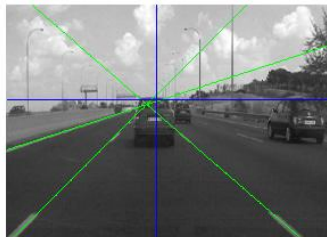


Figura 4.4: Líneas detectadas tras Filtro adaptado

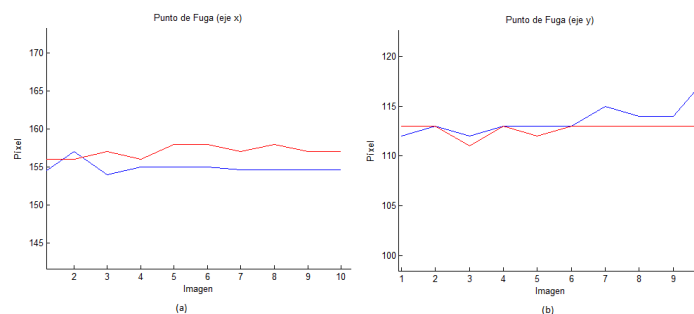


Figura 4.5: Comparación punto de fuga calculado (azul) y punto de fuga teórico (rojo) para el eje-x (a) y el eje-y (b)

### 4.1.2 Parámetros de la determinación de líneas de carretera y punto de fuga

a) Número de máximos de la Transformada de Hough y Filtrado en el ajuste a una senoide.

Podemos delimitar el número de máximos locales que vamos a encontrar en la Transformada de Hough de la imagen y, por lo tanto, el número de rectas que vamos a detectar en la misma.

Es necesario limitar un valor mínimo para garantizar la detección de al menos dos rectas de líneas de carretera y con ellas el punto de fuga. Dado que no podemos garantizar que los valores más votados en la Transformada de Hough se correspondan con las líneas de carretera, es preciso que el número de máximos mínimo sea mayor

de dos. Por otro lado, un valor demasiado elevado de este parámetro ocasiona que se tengan en consideración para la determinación del cálculo del punto de fuga rectas que no se corresponden con las líneas de carretera (Figura 4.6), obteniendo valores del mismo muy distantes del valor teórico. El valor de compromiso, encontrado empíricamente, más adecuado es 3. Los resultados con este valor se muestran en la Figura 4.5 y la Tabla A.2.3, en los que se observan una diferencia teórico-calculado cercana a 0.

Adicionalmente, el sistema nos permite realizar una acotación de las rectas detectadas por Hough para calcular el punto de fuga. Podemos realizar un ajuste a una senoide teniendo en cuenta información a priori para eliminar aquellas rectas que no se corresponden con las líneas de carretera, tal y como explicamos en el apartado 3.3.3.



Figura 4.6: Líneas detectas con número de máximos elevado

Los resultados teniendo en cuenta información adicional, es decir, realizando un filtrado, son notablemente mejores que los que no tienen en cuenta dicha información. Podemos ver la comparativa entre la detección de las rectas en la Figura 4.7. De este modo, los valores obtenidos del punto de fuga distan menos en el caso del filtrado mediante información a priori (Figura 4.8 (b)).

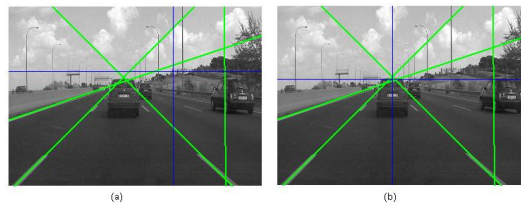


Figura 4.7: Comparativa entre ajuste filtrado (b) y normal (a)

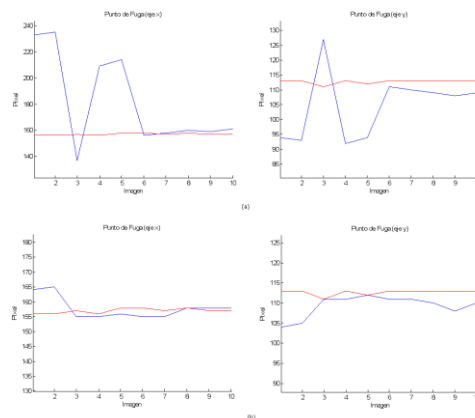


Figura 4.8: Evolución punto fuga calculado (azul) vs punto de fuga teórico (rojo): (a) ajuste normal; (b): ajuste filtrado

d) Filtrado del punto de fuga:  $m$

Como explicamos en el apartado 3.3.4, realizamos un filtrado paso-bajo del punto de fuga para estabilizarlo y eliminar las fluctuaciones. El parámetro que gobierna dicho filtrado es  $m$ , que representa el número de puntos de fuga previos que se consideran para realizar el filtrado.

Elegir un valor de  $m$  muy pequeño no solucionaría las fluctuaciones, puesto que estaríamos realizando un filtrado en el tiempo con una ventana temporal demasiado pequeña, con lo que prácticamente no produciría efecto alguno (Figura 4.9 (a)). Por otro lado, tampoco es aconsejable elegir un valor de  $m$  muy grande, puesto que en ese caso tardaría bastante tiempo en estabilizarse y corregirse esas fluctuaciones (Figura 4.9 (b)). Un valor de  $m$  en torno a 5-10 parece bastante adecuado. En la Figura 4.10 ( $m = 5$ ) observamos cómo el valor filtrado (color azul) se estabiliza y se aproxima mejor al valor teórico (color rojo) a partir de la imagen 5.

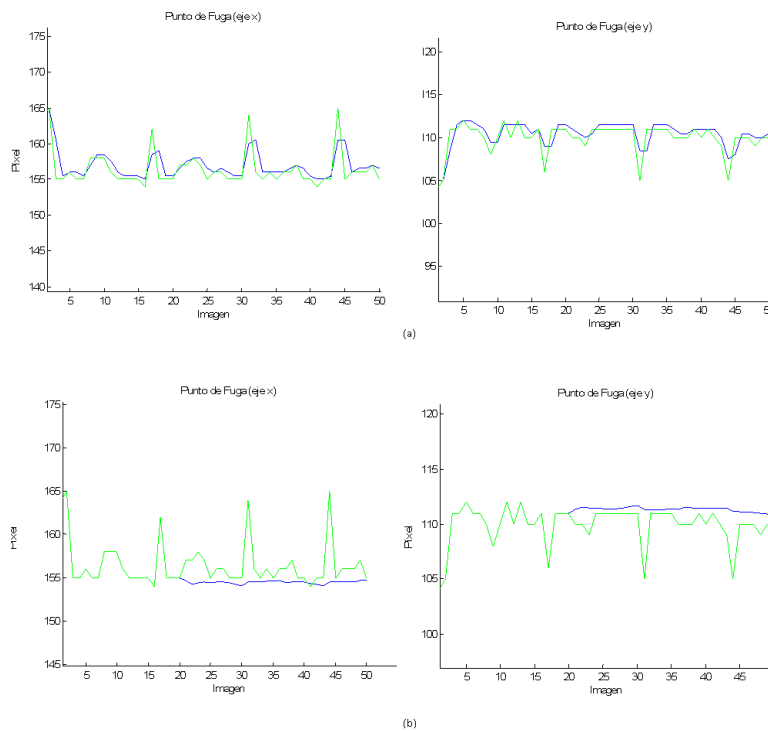


Figura 4.9: Comparación punto fuga calculado (verde) y filtrado (azul): (a) Filtrado con  $m = 2$  ;  
(b) Filtrado con  $m = 20$

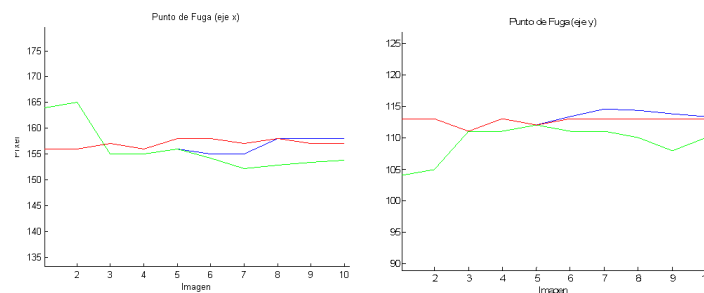


Figura 4.10: Filtrado del punto de fuga: verde: punto de fuga calculado; azul: punto de fuga filtrado; rojo: punto de fuga teórico.

## 4.2 Impacto en la construcción de la IPM:

En este apartado analizamos en detalle, numérica y visualmente, la bondad de la IPM, comparando los resultados obtenidos de ambos métodos de rectificación del plano con nuestro *ground-truth*.

### 4.2.1 Parámetros del Método 1: Homografía con puntos de correspondencia:

a) Gestión oclusiones: simple/avanzada

En los casos en los que se produce una oclusión nos falta información para poder obtener todas las variables necesarias por el sistema. Mediante este parámetro podemos elegir entre proceder con una gestión de oclusiones simple, en la que asumimos que la escena es muy similar entre imágenes consecutivas, con lo que al detectar la oclusión hacemos uso de información del instante anterior, modificándola en base a la variación del punto de fuga. Los resultados realizados con esta alternativa son buenos siempre que la falta de información no se prolongue durante mucho tiempo, pues en tal caso ya no se puede considerar que la imagen mantenga semejanza con la última que presentaba toda la información, ni se produzca un cambio de carril (el movimiento horizontal del vehículo provoca que la imagen anterior no tenga localizadas las líneas de carretera en la misma posición). Al fallar en la obtención de la información de la escena, estamos fallando en la generación de la IPM, produciendo una mala reconstrucción de la escena 3D, con rectas que no son paralelas y con anchuras de carril no constantes. En la Figura 4.11 se observa que en la segunda imagen se produce una oclusión y a partir de la tercera imagen, y debido al cambio de carril que se está realizando, las rectas del instante anterior no sirven como aproximación de las nuevas, provocando IPMs incorrectas: las imágenes rectificadas presentan rectas que no son paralelas, como podemos observar en la Figura 4.11 (b): los círculos rojos marcan aquellas zonas donde se aprecia el hecho de que las rectas no son perfectamente paralelas.

Proceder con la gestión de oclusiones avanzadas nos permite inferir la información necesaria para el proceso a partir de otra información de la escena, tal y como explicamos en 3.4.1. En este caso, los resultados de la IPM son mejores frente a cambios de carril del vehículo. En la Figura 4.12 observamos que, a pesar de producirse una oclusión en la segunda imagen, es capaz de detectar correctamente las líneas de carretera en los instantes posteriores. Esto genera IPMs de mejor calidad: podemos observar en la Figura 4.12 (b) como el rectángulo naranja marca las rectas paralelas.



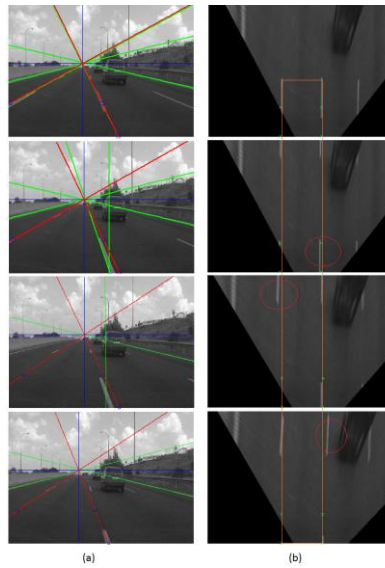


Figura 4.11: Gestión oclusiones simple: (a) Detección rectas en la imagen original; (b): IPMs generadas incorrectamente

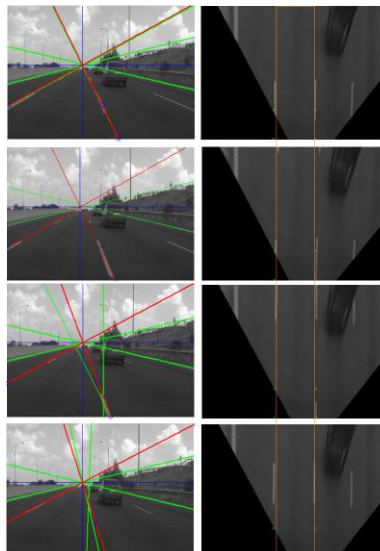


Figura 4.12: Gestión oclusiones avanzada: (a) Detección rectas en la imagen original; (b) IPMs generadas correctamente

#### 4.2.2 Parámetros del Método 2: Rectificación conocida la matriz de proyección:

a)  $Z_{\max}$  :

Este parámetro de diseño determina la profundidad de la IPM a mostrar. Vemos claramente en la Figura 4.13 cuál es su efecto: valores bajos determinan una profundidad baja, es decir, sólo muestra la zona inferior de la imagen (4.13 (a)). Mientras que valores altos determinan gran profundidad (4.13 (b))

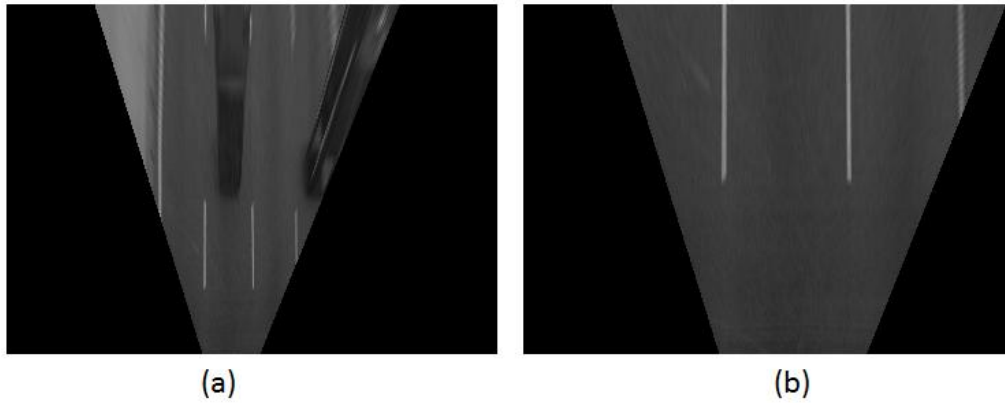


Figura 4.13: Comparativa diferentes profundidades: (a)  $Z_{\max} = 10$  ; (b)  $Z_{\max} = 5$

#### b) Estabilización:

Mediante este parámetro podemos seleccionar si llevar a cabo o no la estabilización de la IPM a partir del punto de fuga. Los resultados mostrados en la Figura 4.14 son contundentes: la IPM estabilizada presenta aspecto correcto, consiguiendo rectas paralelas y de anchura constante.

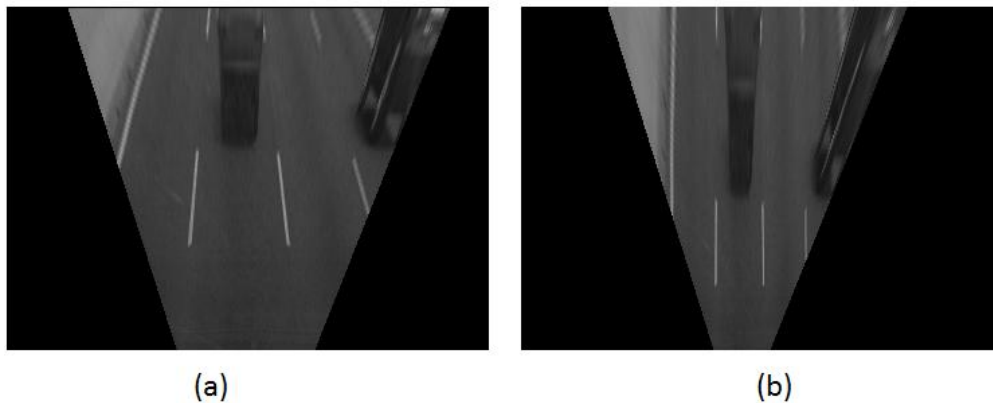


Figura 4.14: Comparativa de IPM no estabilizada (a) y estabilizada (b).

Las razones de que la IPM no estabilizada presente rectas que no son paralelas y anchura de carril variante se debe a que no está teniendo en cuenta las variaciones del sistema de adquisición (los ángulos  $\theta$  y  $\gamma$  no son nulos), tal y como explicamos en el apartado 3.4.2.

### 4.3 Comparación ambos métodos:

El último apartado de la evaluación va a realizar una comparativa entre las dos implementaciones realizadas.

En tramos de la secuencia en los que no se produce movimiento del vehículo ambas implementaciones responden perfectamente. Si observamos la Figura 4.15 ambos métodos generan IPMs en los que las rectas son paralelas (el rectángulo naranja está perfectamente alineado con las líneas de la carretera).

Sin embargo, en los casos en los que se produce algún tipo de maniobra del vehículo (como puede ser un cambio de carril), el método 1 genera rectas que no llegan a ser paralelas del todo (como podemos ver en la Figura 4.16 (b), el rectángulo naranja no se alinea con las líneas de carretera), mientras que el método 2 consigue mantener las rectas paralelas, con anchura del carril constante (Figura 4.16(c)).

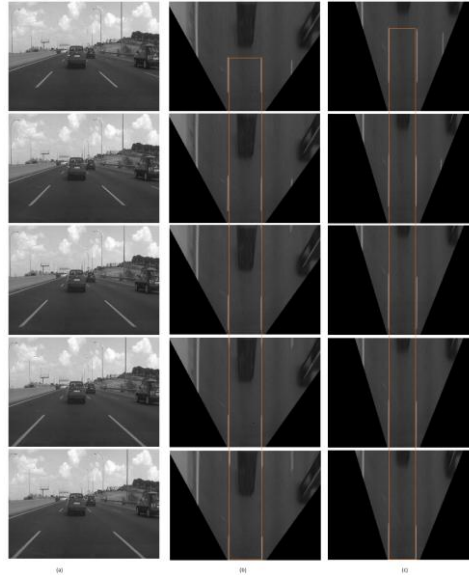


Figura 4.15: Generación IPM en línea recta: (a) Imagen original; (b) IPM Método 1; (c) IPM Método 2

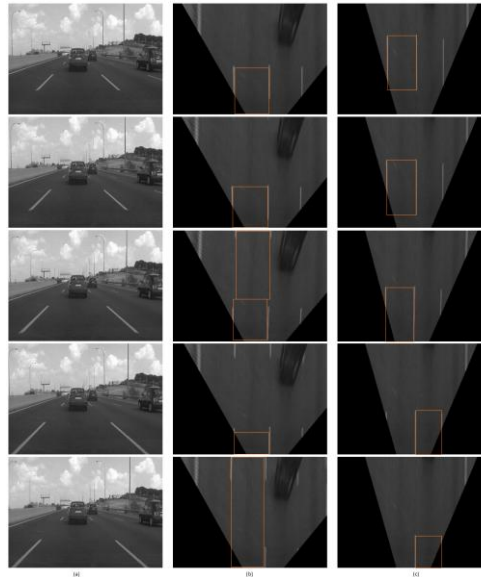


Figura 4.16: Generación IPM en cambio de carril: (a) Imagen original; (b) IPM Método 1; (c) IPM Método 2

Tal y como esperábamos, el método basado en la matriz de proyección proporciona mejores resultados, ya que se basa en un mayor conocimiento a priori de la escena y no depende tanto de la información extraída de la imagen (únicamente necesita estimar el punto de fuga para realizar la estabilización de la IPM).

# 5. Conclusiones y Trabajos Futuros

## 5.1 Conclusiones

El objetivo de este Trabajo Final de Grado era la obtención de imágenes rectificadas de un plano de carretera. Para la consecución de dicho objetivo se han desarrollado dos alternativas diferentes, cada una de ellas con sus ventajas e inconvenientes.

Como se explicó anteriormente, la determinación del punto de fuga dominante de la imagen es un paso clave para ambas implementaciones. Una vez analizados los resultados se concluye que la técnica empleada para la determinación del mismo, mediante un sistema de votaciones (la transformada de Hough), no siempre proporciona unos resultados idóneos. No obstante, los diferentes filtrados (temporales y espaciales) realizados sobre el proceso de determinación del punto de fuga consiguen mejorar los resultados obtenidos.

Las imágenes rectificadas obtenidas por ambos métodos son exitosas para diferentes aplicaciones de sistemas avanzados de ayuda a la conducción (ADAS) que no necesitan una rectificación métrica, sino simplemente medidas relativas, tales como la posición trasversal del vehículo en el carril propio, la velocidad relativa con respecto a otros vehículos, distancias entre vehículos, etc.

Comparando ambos métodos de rectificación, queda claramente de manifiesto que los mejores resultados se consiguen con el segundo de ellos, aquél que emplea el conocimiento de la matriz de proyección. Este método no precisa de información de la escena para el cómputo de la IPM, y a cambio necesita mucho conocimiento a priori: la matriz de calibración. Este caso hace la asunción de que el ángulo de alabeo es nulo (algo típicamente asumible siempre que la instalación de la cámara haya sido realizada correctamente). Adicionalmente, permite una estabilización de la IPM, teniendo en cuenta para ello el punto de fuga de la escena. Con la estabilización conseguimos mantener IPMs correctas en situaciones en las que los ángulos de rotación no son nulos (cambios de pendiente y cambios de carril).

Opuesto al método 2, el primer método propuesto no necesita de ninguna información previa sobre la escena con la que estamos trabajando. No obstante, para poder conseguir la rectificación de la imagen mediante una homografía, precisa de varias asunciones: existen líneas de carretera y se encuentran en un plano (plano de la carretera). Además dichas líneas se encuentran perfectamente alineadas con respecto al eje-Z. Este método es muy sensible a cambios de pendiente y de carril. Además, es necesario llevar a cabo una gestión de oclusiones en los que no se consigue extraer la información necesaria. Dichas oclusiones son debidas a la presencia de otros vehículos, cambios de iluminación, cambios meteorológicos, etc.

En definitiva, a pesar de la simplicidad del método 1, es tan sensible a variaciones en la escena, que los resultados obtenidos no son lo suficientemente

buenos. Mientras que el método 2, es capaz de generar IPMs muy fiables casi frente a cualquier situación.

## 5.2 Trabajos futuros

Una vez vistas los resultados de nuestro sistema y analizadas las limitaciones del mismo, se proponen una serie de trabajos futuros para, manteniendo la misma línea de investigación, conseguir mejoras en los resultados y rendimiento del sistema:

- Relativo a la localización de las líneas de carretera:
  - Realizar un análisis en mayor profundidad (zonas de interés) sobre la Transformada de Hough con el objetivo de tener una mejor estimación de las rectas correspondientes a las líneas de carretera.
- En relación con la determinación del punto de fuga, se proponen varias alternativas para mejorar los resultados obtenidos:
  - Mejorar el sistema implementado, mediante filtros de Kalman, para conseguir mejores estabilizaciones del punto de fuga.
  - Determinar varios puntos de fuga
  - Emplear aproximaciones más complejas que introducen modelos de error, definiendo funciones de coste que hay que minimizar.
- Con respecto al método 1:
  - Emplear otras técnicas para la detección de los puntos de correspondencia, sin asumir que las líneas tienen que estar perfectamente alineadas con respecto al eje-Z, consiguiendo así adaptarnos a cambios de carril.
- Relacionado con el método 2:
  - Obtener la matriz de calibración real de forma exacta
  - Mejoras relacionadas con la determinación del punto de fuga
- Implementar otras alternativas para la rectificación de planos de imagen, propuestos en el apartado 2.4.
- Realizar implementaciones en lenguajes de programación de nivel medio (C, C++), más eficientes computacionalmente, con el fin de llegar al objetivo de trabajar en tiempo real.

# Referencias

- [1] M. Nieto and L. Salgado. Real-time robust estimation of vanishing points through nonlinear optimization. In *SPIE Proc. International Conference on Real-Time Image and Video Processing*, page (Invited Paper), 2010.
- [2] E. Lutton, H. Maître, and J. Lopez-Krahe. Contribution to the determination of vanishing points using Hough transform. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 16(4): 430-438, 1994.
- [3] N. Simond and P. Rives. Homography from a vanishing point in urban scenes. In *IEEE/RSJ Proc. International Conference on Intelligent Robots and Systems*, volume 1, pages 1005-1010, 2003.
- [4] A.M. Muad, A.Hussain, S.A. Samad, M.M. Mustaffa, and B.Y. Majlis. Implementation of inverse perspective mapping algorithm for the development of an automatic lane tracking system. In *IEEE Proc. Region 10 Conference (TEN-CON)*, volume 1, pages 207-211, 2004.
- [5] T. Suttorp and T. Bücher. Robust vanishing point estimation for driver assistance. In *IEEE Proc. Intelligent Transportation Systems Conference*, pages 1550-1555, 2006.
- [6] M. Nieto, L. Salgado, F. Jaureguizar, and J. Cabrera. Stabilization of inverse perspective mapping images based on robust vanishing point estimation. In *IEEE Proc. Intelligent Vehicles Symposium 2007*, pages 315-320, 2007.
- [7] D. Alonso, L. Salgado, and M. Nieto. Robust vehicle detection through multidimensional classification for on board video based. In *IEEE Proc. International Conference on Image Processing*, pages 321-324, 2007.
- [8] M. Nieto, J. Arróspide, L. Salgado, and F. Jaureguizar. On-board video based system for robust road modeling. In *IEEE Proc. Content-Based Multimedia Indexing*, pages 109-116, 2008.
- [9] C. Maduro, K. Batista, P. Peixoto, and J. Batista. Estimation of vehicle velocity and traffic intensity using rectified images. In *IEEE International Conference on Image Processing*, pages 777-780, 2008.
- [10] J. Arróspide, L. Salgado, M. Nieto, and F. Jaureguizar. On-board robust multiple vehicle detection and tracking using adaptive quality evaluation. In *IEEE Proc. International Conference on Image Processing*, pages 2008-2011, 2008.

- [11] M. Nieto, J. Arróspide, L. Salgado, and F. Jaureguizar. Robust multiple lane road modeling based on perspective analysis. In *IEEE Proc. International Conference on Image Processing*, pages 2396-2399, 2008.
- [12] J. Arróspide, M. Nieto, L. Salgado, and R. Mohedano. Homography-based ground plane detection using a single on-board camera. *IET Intelligent Transportation Systems*, page (accepted), 2010.
- [13] J. Arróspide, L. Salgado, and M. Nieto. Vehicle detection and tracking using homography-based plane rectification and particle filtering. In *IEEE Proc. Intelligent Vehicles Symposium*, page (accepted), 2010.
- [14] M. Nieto and L. Salgado. Plane rectification through robust vanishing point tracking using the Expectation-Maximization algorithm. In *IEEE Proc. International Conference on Image Processing*, page (accepted), 2010.
- [15] M. Nieto, J. Arróspide, and L. Salgado. Road environment modeling using robust perspective analysis and recursive Bayesian segmentation. *Machine Vision and Applications*, page (accepted), 2010.
- [16] M. Bertozzi, and A. Broggi. Gold: A parallel real-time stereo vision system for generic obstacle and lane detection. In *IEEE Trans. on Image Processing*, volume 7, No 1, pages 62-81, 1998.
- [17] J. P. González, and U. Ozguner. Lane detection using histogram-based segmentation and decision tree. In *IEEE Proc. Int. Trans. Systems*, pages 346-351, 2000.
- [18] S. Beucher, and X. Yu. Road recognition in complex traffic situations. In *7<sup>th</sup> Symp. on Trans. Systems*, pages 423-428, Tianjin, China, August 24-26, 1994.
- [19] K. Macek, B. Williams, S. Kolski, and R. Siegwart. A lane detection vision module for driver assistance. In *IEEE/APS Proc. Conf. on Mech. and Robotics*, Aachen, Germany, 2004.
- [20] S. Nedevschi, R. Schmidt, T. Graf, R. Danescu, D. Frentiu, T. Marita, F. Oniga, and C. Pocol. 3D Lane Detection System Based on Stereovision. In *IEEE Proc. Intelligent Transportation Systems Conference*, pages 161-166, 2004.
- [21] S. Nedevschi, F. Oniga, and R. Danescu. Increased Accuracy Stereo Approach for 3D Lane Detection. In *IEEE Proc. Intelligent Vehicule Symposium*, pages 42-49, 2006.
- [22] R. Danescu, S. Nedevschi, and M. Meinecke. A Stereovision-Based Probabilistic Lane Tracker for Difficult Road Sceneraios. In *IEEE Proc. Intelligent Vehicles Symposium*, pages 536-541, 2008.

- [23] A. Broggi, M. Bertozzi, and A. Fascioli. Self-calibration of a Stereo Vision System for Automotive Applications. In *IEEE Proc. Int. Conf. on Robotics and Automation*, volume 4, pages 3698-3703, 2001.
- [24] P. Chang, D. Hirvonen, T. Camus, and B. Southall. Stereo-Based Object Detection, Classification, and Quantitative Evaluation with Automotive Applications. In *IEEE Proc. Computer Vision and pattern Recognition*, volume 3, pages 62-68, 2005.
- [25] J. Zhou, and B. Li: Homography-based Ground Detection for A Mobile Robot Platform Using a Single Camera. In *IEEE International Conference on Robotics and Automation*, pages 4100-4105, 2006.
- [26] Marcos Nieto. Detection and tracking of vanishing points in dynamic environments. In Universidad Politécnica de Madrid, Ph. D. Thesis.
- [27] J. Bescós. Diapositivas del tema 4 de la asignatura: Tratamiento de Señales Visuales, 2014.
- [28] M. Bertozzi, and A. Broggi. A parallel real-time stereo vision system for generic obstacle and lane detection. In *IEEE Trans. Image Process.*, 7(1), pages 62-81, 1998.
- [29] Hartley, R. I. and Zisserman, A., *Multiple View Geometry in Computer Vision*, Cambridge University Press, June 1999.





# Anexos

## A.1 Algoritmo DLT:

El algoritmo DLT (*Direct Linear Transformation*) computa la ecuación  $X' = HX$  en términos del producto vectorial como  $X'_i \times (HX_i) = 0$ , siendo  $X'_i = (x'_i, y'_i, z'_i)$  y  $X_i = (x_i, y_i, z_i)$ .

Desarrollando  $(HX_i)$ :

$$HX_i = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} \begin{bmatrix} x_i \\ y_i \\ z_i \end{bmatrix} = \begin{bmatrix} h_{11}x_i + h_{12}y_i + h_{13}z_i \\ h_{21}x_i + h_{22}y_i + h_{23}z_i \\ h_{31}x_i + h_{32}y_i + h_{33}z_i \end{bmatrix}$$

En segundo lugar, realizamos el producto vectorial:

$$\begin{aligned} X'_i \times (HX_i) &= \begin{bmatrix} x'_i \\ y'_i \\ z'_i \end{bmatrix} \times \begin{bmatrix} h_{11}x_i + h_{12}y_i + h_{13}z_i \\ h_{21}x_i + h_{22}y_i + h_{23}z_i \\ h_{31}x_i + h_{32}y_i + h_{33}z_i \end{bmatrix} = \\ &= \begin{bmatrix} i & j & k \\ x'_i & y'_i & z'_i \\ h_{11}x_i + h_{12}y_i + h_{13}z_i & h_{21}x_i + h_{22}y_i + h_{23}z_i & h_{31}x_i + h_{32}y_i + h_{33}z_i \end{bmatrix} = \\ &= i[(h_{31}x_i + h_{32}y_i + h_{33}z_i)y'_i - (h_{21}x_i + h_{22}y_i + h_{23}z_i)z'_i] + \\ &\quad j[(h_{11}x_i + h_{12}y_i + h_{13}z_i)z'_i - (h_{31}x_i + h_{32}y_i + h_{33}z_i)x'_i] + \\ &\quad k[(h_{21}x_i + h_{22}y_i + h_{23}z_i)x'_i - (h_{11}x_i + h_{12}y_i + h_{13}z_i)y'_i] \end{aligned}$$

De forma matricial quedaría:

$$\begin{bmatrix} 0 & 0 & 0 & -(z'_i x_i) & -(z'_i y_i) & -(z'_i z_i) & (y'_i x_i) & (y'_i y_i) & (y'_i z_i) \\ (z'_i x_i) & (z'_i y_i) & (z'_i z_i) & 0 & 0 & 0 & -(x'_i x_i) & -(x'_i y_i) & -(x'_i z_i) \\ -(y'_i x_i) & -(y'_i y_i) & -(y'_i z_i) & (x'_i x_i) & (x'_i y_i) & (x'_i z_i) & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} h_{11} \\ h_{12} \\ h_{13} \\ h_{21} \\ h_{22} \\ h_{23} \\ h_{31} \\ h_{32} \\ h_{33} \end{bmatrix} = 0$$

Tiene la forma  $A_i h = 0$ , donde  $A_i$  es una matriz 3x9 y  $h$  es un vector columna 9x1 formado por los parámetros de la matriz  $H$ .

Hay que tener en cuenta que los parámetros  $z'_i$  y  $z_i$  de la matriz  $A_i$  son 1, puesto que son coordenadas de la imagen. Además, la 3ª fila de la matriz  $A_i$  no es linealmente independiente. Por ello la matriz  $A_i$  queda simplificada a una matriz de 2x9:

$$\begin{bmatrix} 0 & 0 & 0 & -(z'_i x_i) & -(z'_i y_i) & -(z'_i z_i) & (y'_i x_i) & (y'_i y_i) & (y'_i z_i) \\ (z'_i x_i) & (z'_i y_i) & (z'_i z_i) & 0 & 0 & 0 & -(x'_i x_i) & -(x'_i y_i) & -(x'_i z_i) \end{bmatrix}$$

Como estamos en un caso de 4 puntos de correspondencia, tendremos 4 matrices  $A_i$ . Juntándolas, formaremos la matriz  $A$  de tamaño 8x9:

$$A = \begin{bmatrix} 0 & 0 & 0 & -(z'_1 x_1) & -(z'_1 y_1) & -(z'_1 z_1) & (y'_1 x_1) & (y'_1 y_1) & (y'_1 z_1) \\ (z'_1 x_1) & (z'_1 y_1) & (z'_1 z_1) & 0 & 0 & 0 & -(x'_1 x_1) & -(x'_1 y_1) & -(x'_1 z_1) \\ 0 & 0 & 0 & -(z'_2 x_2) & -(z'_2 y_2) & -(z'_2 z_2) & (y'_2 x_2) & (y'_2 y_2) & (y'_2 z_2) \\ (z'_2 x_2) & (z'_2 y_2) & (z'_2 z_2) & 0 & 0 & 0 & -(x'_2 x_2) & -(x'_2 y_2) & -(x'_2 z_2) \\ 0 & 0 & 0 & -(z'_3 x_3) & -(z'_3 y_3) & -(z'_3 z_3) & (y'_3 x_3) & (y'_3 y_3) & (y'_3 z_3) \\ (z'_3 x_3) & (z'_3 y_3) & (z'_3 z_3) & 0 & 0 & 0 & -(x'_3 x_3) & -(x'_3 y_3) & -(x'_3 z_3) \\ 0 & 0 & 0 & -(z'_4 x_4) & -(z'_4 y_4) & -(z'_4 z_4) & (y'_4 x_4) & (y'_4 y_4) & (y'_4 z_4) \\ (z'_4 x_4) & (z'_4 y_4) & (z'_4 z_4) & 0 & 0 & 0 & -(x'_4 x_4) & -(x'_4 y_4) & -(x'_4 z_4) \end{bmatrix}$$

## A.2 Tablas:

Tabla A.2.1: Valores teórico-calculado del punto de fuga para Canny, mejor situación:

Imagen	Puto de fuga calculado		Punto de fuga teórico		Distancia (diferencia)		Distancia (módulo)
	eje-x	eje-y	eje-x	eje-y	eje-x	eje-y	
1	156	111	156	113	0	2	2
2	157	111	156	113	1	2	2,236067977
3	156	111	157	111	1	0	1
4	157	113	156	113	1	0	1
5	156	113	158	112	2	1	2,236067977
6	157	113,4	158	113	1	0,4	1,077032961
7	156	113,8	157	113	1	0,8	1,280624847
8	158	114	158	113	0	1	1
9	158	113,8	157	113	1	0,8	1,280624847
10	157	113,6	157	113	0	0,6	0,6
MEDIA					0,8	0,86	1,371041861

Tabla A.2.2: Valores teórico-calculado del punto de fuga para Canny, mala situación:

Imagen	Puto de fuga calculado		Punto de fuga teórico		Distancia (diferencia)		Distancia (módulo)
	eje-x	eje-y	eje-x	eje-y	eje-x	eje-y	
1	164	104	156	113	8	9	12,04159458
2	165	105	156	113	9	8	12,04159458
3	155	111	157	111	2	0	2
4	155	111	156	113	1	2	2,236067977
5	156	112	158	112	2	0	2
6	155	111	158	113	3	2	3,605551275
7	155	111	157	113	2	2	2,828427125
8	158	110	158	113	0	3	3
9	158	108	157	113	1	5	5,099019514
10	158	110	157	113	1	3	3,16227766
MEDIA					2,9	3,4	4,801453271

Tabla A.2.3: Valores teórico-calculado del punto de fuga para Filtrado adaptado:

Imagen	Puto de fuga calculado		Punto de fuga teórico		Distancia (diferencia)		Distancia (módulo)
	eje-x	eje-y	eje-x	eje-y	eje-x	eje-y	
1	154	112	156	113	2	1	2,236067977
2	157	113	156	113	1	0	1
3	154	112	157	111	3	1	3,16227766
4	155	113	156	113	1	0	1
5	155	113	158	112	3	1	3,16227766
6	155	113	158	113	3	0	3
7	154,6	115	157	113	2,4	2	3,12409987
8	154,6	114	158	113	3,4	1	3,544009029
9	154,6	114	157	113	2,4	1	2,6
10	154,6	118	157	113	2,4	5	5,54616985
MEDIA					2,36	1,2	2,837490205